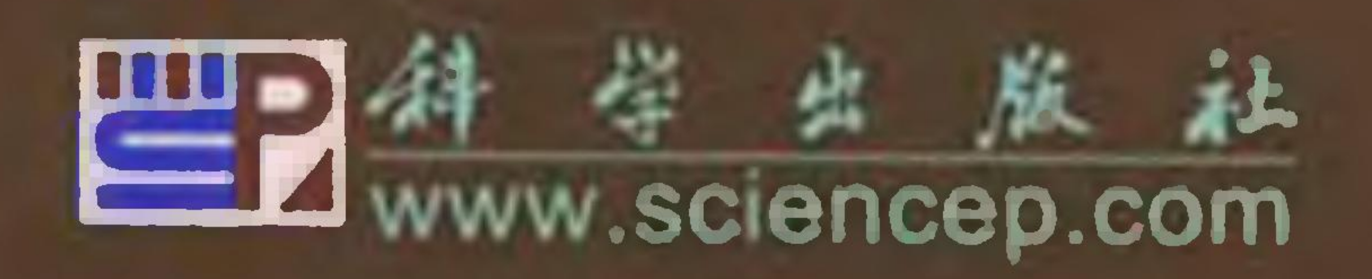
第一卷　第二分册

数学名著译丛

微积分和
数学分析引论

R.柯朗　F.约翰　著



目　　录

第五章 泰勒展开式 ……………………………… (463)

第四章　在物理和几何中的应用

4.1　平面曲线理论

a. 参数表示

定义　用方程 $y=f(x)$ 来表示曲线在几何上有很大的限制：这样表示的曲线与平行于 y 轴的任意直线相交不能多于一点．通常，把曲线分成可以表为 $y=f(x)$ 的若干部分来克服这个限制．因此，一个以原点为中心，以 a 为半径的圆，可由定义于 $-a\leqslant x\leqslant a$ 的两个函数 $y=\sqrt{a^2-x^2}$ 和 $y=-\sqrt{a^2-x^2}$ 给出．但是，对于像平行于 y 轴的直线，这个办法却不行．

表示曲线的更灵活的方法是利用方程 $\varphi(x, y)=0$ 这样一种隐示表示法，在这个方程里包含一个含有两个自变量的函数 φ．例如，以原点为中心，以 a 为半径的圆可由 $\varphi(x, y)=x^2+y^2-a^2=0$ 完全描述．平面上任意直线都有一个形如 $ax+by+c=0$ 的隐示方程，其中 a, b, c 是常数，且 a 和 b 不都是零．在 $b=0$ 时，我们得到 y 轴的一条平行线．

曲线的这种隐示描述法有一个缺点，即欲求曲线上的点 (x, y)，就必须对于一给定的 x，求解方程 $\phi(x, y)=0$．这个问题我们将在第二卷详细讨论．

曲线最直接和最灵活的描述法是参数表示．我们不把直角坐标 x 或 y 中的一个看成另一个的函数，而把两个坐标 x 和 y 都看成为第三个自变量 t 的函数，t 是所谓参数或参量 (parameter)[1]；当 t 在一个相应的区间内变化时，坐标为 x 和 y 的点就描绘出一条曲线，这样的参数表示我们已经遇到过了；例如，圆 $x^2+y^2=a^2$ 有参数表示 $x=a\cos t,\ y=a\sin t$．这里，t 表示圆的中心角．

---

1) 这个词表示辅助变量而不是主要变量．

对于椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$，我们有类似的参数表示 $x=a\cos t$，$y=b\sin t$，其中 t 是所谓偏心角，即由椭圆上的点 $P=(a\cos t, b\sin t)$ 向上或向下引垂线与外接圆相交的交点的圆心角. 这里，我们假定 $b<a$（见图 4.1）. 在上述两例中，当 t 在区间 $0\leqslant t<2\pi$ 内变化时，坐标为 x，y 的点描绘了整个的圆或椭圆.

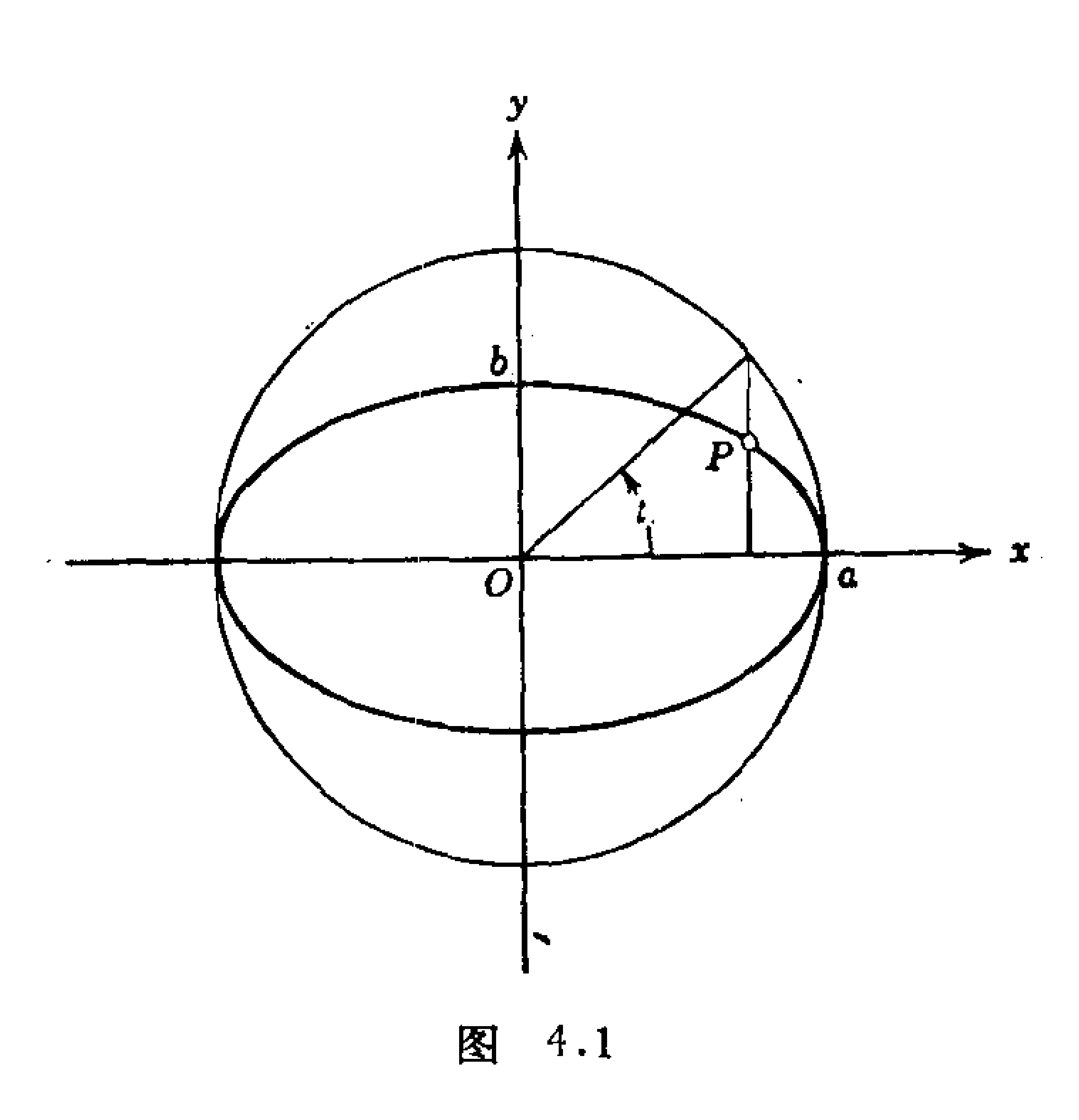


图 4.1

一般地，曲线 C 表示为参数 t 的两个函数，

$$x=\phi(t)=x(t),$$
$$y=\psi(t)=y(t);$$

在不会发生混淆的时候，我们将使用符号 $x(t)$ 和 $y(t)$.[1)]

除非特别指明，我们总是假定 ϕ 和 ψ 有连续导数.

参数区间到曲线上的映射；方向的指向

对于给定的曲线，必须这样确定两个函数 $\phi(t)$ 和 $\psi(t)$，使得对应于 t 的某个区间，一对函数值 $x(t)$ 和 $y(t)$ 的集合定义了曲线上的全部的点，而且没有其他点. 这样，在曲线上的点和 t 轴的一个区间上的 t 值之间就有了一个对应关系. 参数表示定义了 t 轴到曲线的一个映射，即由 t 轴上的原象点 t 映射到 C 的点 $x=\phi(t)$，$y=\psi(t)$.

因为假定 $x(t)$ 和 $y(t)$ 是连续的，所以 t 轴上的相邻点对应到曲线上的相邻点. 因为 t 轴的点是有序的，所以我们可以用明显的方式把 C 的点规定一个次序或"指向"，即如果 $t_1<t_2$，那么由 t_1

1) 符号 $x=\phi(t)$ 等，侧重于表示自变量和因变量之间的特殊函数关系；而符号 $x(t)$ 等，是指，把 t 看成自变量，x 是它的函数.

映射的点在由 t_2 映射的点的前面(见 358 页). 因此，对于曲线这个含糊的直观概念，参数表示给出了明确的意义，把它作为点的集合，其中的点象按直线上的顺序一样排列起来了.

b. 参数变换

参数 t 的值可用来区别曲线 C 上的不同的点；它们对曲线上的每个点起了“命名”的作用.

同一曲线 C 可以有许多不同的参数表示. 沿着曲线连续变化并且在曲线不同的点上取不同值的任何一个量都可以当作参数.

例如，如果曲线原来由方程 $y=f(t)$ 给出，那么我们可以选择变量 x 作为参数 t. 而用函数 $x=t$, $y=f(t)$ 描述曲线. 类似地，对于给定 x 作为 y 的函数所描述的曲线，比如说 $x=g(y)$，我们可以用 y 作为参数 t，记作 $x=g(t)$，$y=t$.

在极坐标 r，θ 中，用方程 $r=h(\theta)$ 给定的曲线(见第一章，第 108 页)我们可以选择 θ 作为参数 t 而得到参数表示

$$x=r\cos\theta=h(t)\cos t=\phi(t),$$

$$y=r\sin\theta=h(t)\sin t=\psi(t).$$

从一给定的曲线 C 的参数表示 $x=\phi(t)$，$y=\psi(t)$，我们总可以推导出很多其他的参数表示. 为此目的，我们取任意一个函数 $\tau=\chi(t)$，它在相应于曲线 C 的点的 t 区间上单调并且连续；那么在相应的 τ 区间上，函数 χ 存在一个单调和连续的反函数 $t=\sigma(\tau)$. 这时 C 的点 (x,y) 的坐标可表为

$$x=\phi[\sigma(\tau)]=\alpha(\tau),\quad y=\psi[\sigma(\tau)]=\beta(\tau).$$

函数 $\alpha(\tau)$ 和 $\beta(\tau)$ 也是连续的；并且 C 的不同点对应于 t 的不同值，因此，由于函数 σ 的单调性也就对应于不同的 τ 值. 参数由 t 到 τ 的变换的全部效果就在于把 C 上的点“再命名”.

例如，直线 $y=x$ 有参数表示 $x=t$，$y=t$，其中 $-\infty<t<\infty$. 作替换 $\tau=t^3$ 便对这同一条直线给出参数表示 $x=\tau^{1/3}$，$y=\tau^{1/3}$.

类似地，椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 可以有参数表示 $x=a\cos t$，$y=b\sin t$，其中 $0\leqslant t<2\pi$. 定义 $t=c\zeta+d$，c 与 d 为实数 $(c\neq 0)$，同一个椭圆就得到另一个表示 $x(\zeta)=a\cos(c\zeta+d)$，$y(\zeta)=b\sin(c\zeta+d)$；在 $c>0$ 时，ζ 在区间 $-\frac{d}{c}\leqslant\zeta<\frac{2\pi-d}{c}$ 上变化；在 $c<0$ 时，ζ 在区间 $\frac{(2\pi-d)}{c}<\zeta\leqslant-\frac{d}{c}$ 上变化. 作替换 $\tau=\tan\left(\frac{t}{2}\right)$ 可以导出椭圆的"有理"参数表示（见第 314 页）

$$x=\frac{a(1-\tau^2)}{1+\tau^2},\quad y=\frac{2b\tau}{1+\tau^2};$$

当 τ 取遍全部实数值时，我们就得到了这一椭圆所有的点，只缺少点 $S=(-a,0)$.

如果使用适当的参数，可使通常表示式的奇异性消失. 例如，我们可用光滑的函数 $x=t^3$，$y=t^2$ 表示曲线 $y=\sqrt[3]{x^2}$. 当 t 由 $-\infty$ 变到 $+\infty$ 时，坐标为 x, y 的点就描绘出整个曲线（半立方抛物线）.

这种参数选择的灵活性常常使我们能够简化几何性质的研究，当然，几何性质并不依赖于特殊的表示法.

特殊地，有时我们发现使用 $y=f(x)$ 表示 C 或 C 的一部分是方便的. 如果对于曲线的一部分 $t_0\leqslant t\leqslant t_1$，函数 ϕ，ψ 中的一个，譬如说 $x=\phi(t)$ 是单调的，这样的表示法总是可能的. 事实上，对于这部分有唯一的反函数 $t=\gamma(x)$，因此 $y=\psi[\gamma(x)]$.[1)]

c. 沿曲线的运动. 时间作为参量. 摆线的例子

沿曲线的运动.

参量 t 常常有自然的物理意义，即时间. 在平面上一个点的

1）当然，这只是关于曲线在"小范围"性质的一个说法. 就是说，这个说法仅仅对适当小的一部分成立. 通常（例如，在圆的情况下），变量 x 不能在整个曲线上，而只能在一部分曲线上用作参量.

任何运动都可以把它的坐标 x 和 y 表示为时间的函数，在时刻 t，点 (x, y) 在 $(x(t), y(t))$. 这两个函数以参数形式确定了沿路径(即轨道) C 的运动. 它们构成了时间标度到轨道的一个映射.[1]

摆线和次摆线

以摆线为例，当一个圆沿着一条直线或另一个圆匀速而无滑动地滚动时，动圆上一点的路径叫做摆线. 最简单的情况是一个半径为 a 的圆沿 x 轴滚动；圆周上点 P 的路径就是"普通"摆线. 我们这样选择坐标系的原点和初始时间，使得在时刻 $t=0$，点 P 位于原点，在时刻 t 时，圆从它的初始位置转了一个角 t. 这就是说，圆以角速度 1 按顺时针转动. 已经假定圆沿 x 轴作没有滑动地匀速滚动，因此在时刻 t，切点和原点的距离刚好等于由切点到 P 点的弧长. 因此，在时刻 t，滚动圆的中心 M 必定在点 (at, a) 上；圆心以常速 a 向右运动(见图 4.2). 对于在时刻 t 时 P 的坐标，我们得到参数表示

$$x = a(t - \sin t), \quad y = a(1 - \cos t). \tag{1}$$

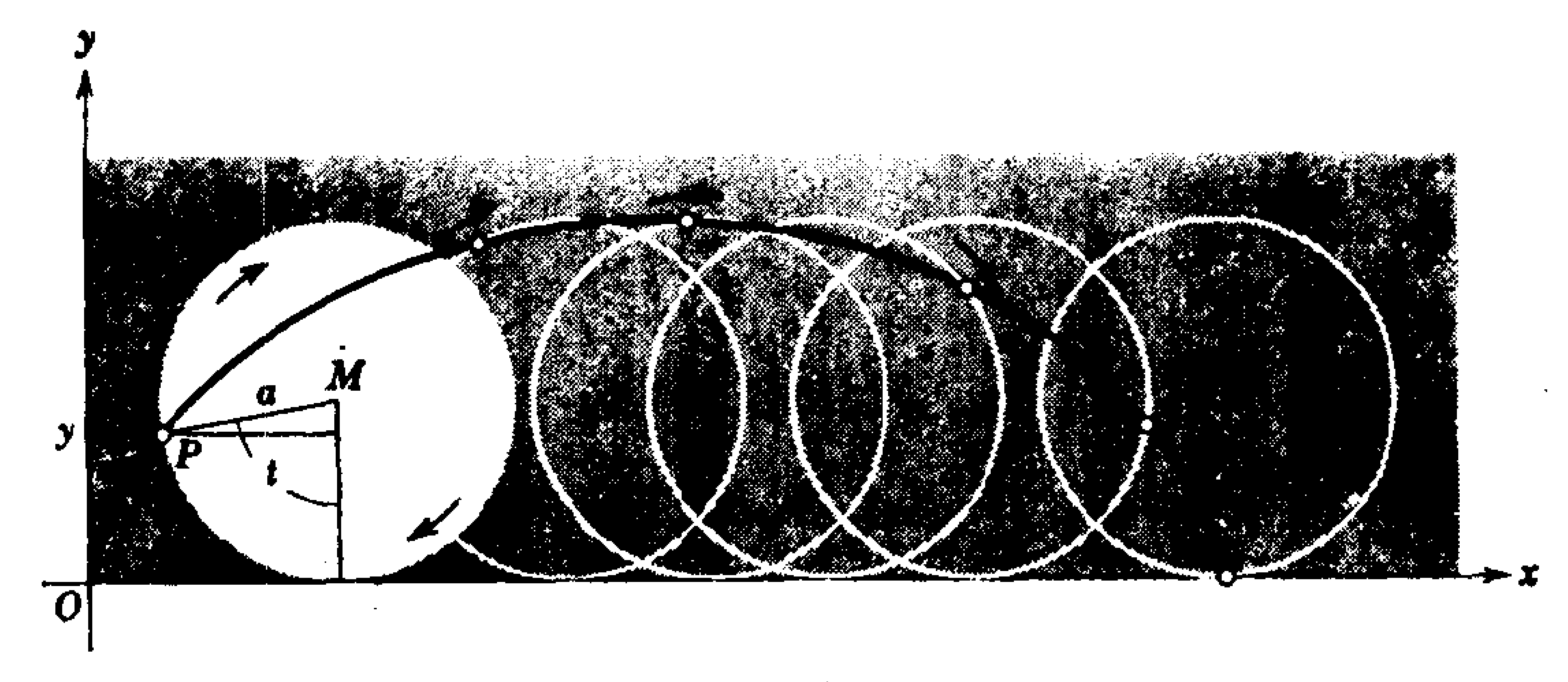


图 4.2 摆线

消去参数 t，我们能得到非参数形式的曲线方程，然而却失去了表达的简洁性，我们有

$$\cos t = \frac{a-y}{a}, \quad t = \arccos\frac{a-y}{a}, \quad \sin t = \pm\sqrt{1-\frac{(a-y)^2}{a^2}},$$

1) 参数 t 的变换对应于时间标度的变换，根据它，曲线 C 用动点描述.

因此

$$x = a\arccos\frac{a-y}{a} \mp \sqrt{y(2a-y)}, \qquad (1a)$$

这就得到 x 作为 y 的函数.

外摆线

我们的下一个例子是外摆线，它定义为一个半径为 c 的圆沿着第二个半径为 a 的圆的圆周外边匀速滚动时，固定在动圆圆周上的点 P 的路径. 设固定的圆以 x，y 平面的原点为中心. 假定动圆沿着固定的圆以这样的方式滚动，即在时刻 t，动圆的中心绕原点转了大小为 t 的角 (图 4.3). 那么，对于在时刻 t，点 $P=(x(t), y(t))$ 的位置 (它在时刻 $t=0$ 是切点 $(a, 0)$)，我们得到参数方程

$$\begin{aligned} x(t) &= (a+c)\cos t - c\cos\left(\frac{a+c}{c}t\right), \\ y(t) &= (a+c)\sin t - c\sin\left(\frac{a+c}{c}t\right). \end{aligned} \qquad (2)$$

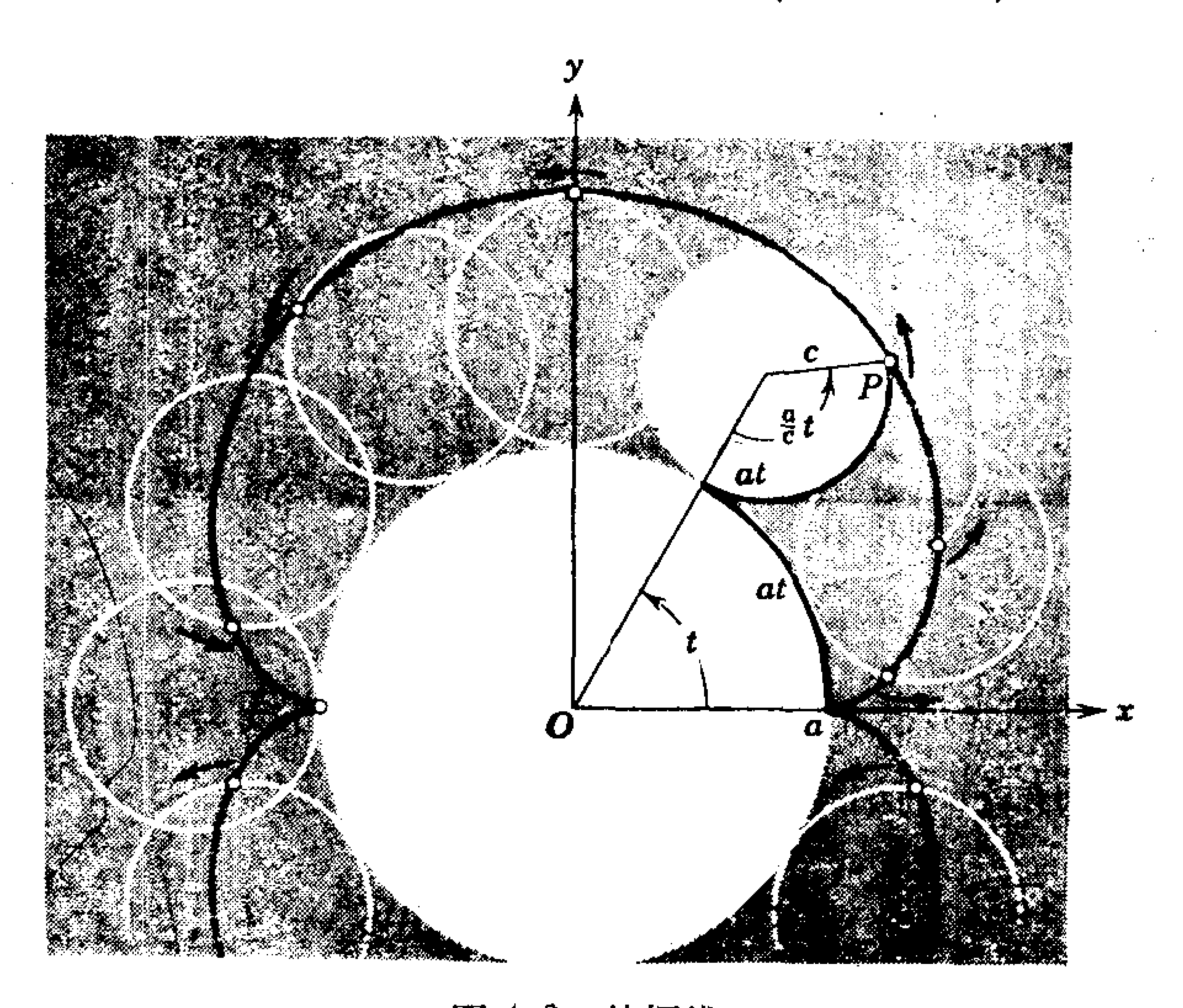


图 4.3 外摆线

当 $a=c$ 时，形成的曲线叫做心脏线 (图 4.4)，参数方程为

$$x(t) = 2a\cos t - a\cos(2t),$$
$$y(t) = 2a\sin t - a\sin(2t). \tag{3}$$

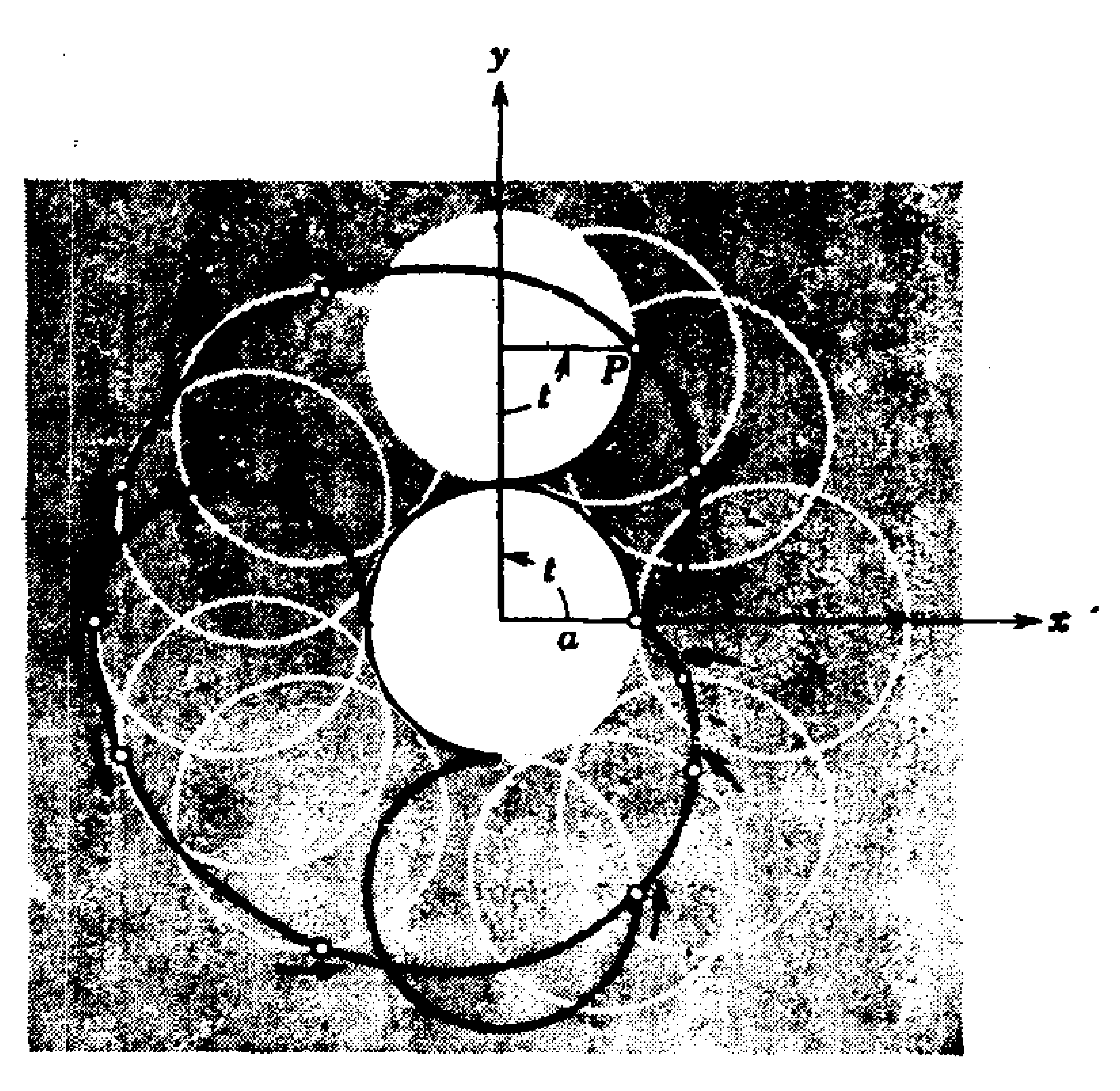


图 4.4　心脏线

第三种摆线是当一个圆沿着另一个固定圆的圆周在内部滚动时，动圆圆周上一点的轨迹. 为了得到这个"内摆线"的参数方程，设固定圆的半径为 a，滚动圆的半径为 c. 设动圆的圆周上点 P 在时刻 $t=0$ 位于 $(a, 0)$. 再假定滚动的圆沿着固定的圆以这样的方式滚动，即在时刻 t，该圆的中心绕原点转了大小为 t 的角度（图 4.5）. 这样，我们得到内摆线的参数方程为

$$x(t) = (a-c)\cos t + c\cos\left(\frac{a-c}{c}t\right),$$
$$y(t) = (a-c)\sin t - c\sin\left(\frac{a-c}{c}t\right). \tag{4}$$

在特殊情况下，当固定圆的半径为动圆半径的 2 倍时，即 $c=\frac{a}{2}$，我们有

$$x(t) = a\cos t,$$
$$y(t) = 0,$$

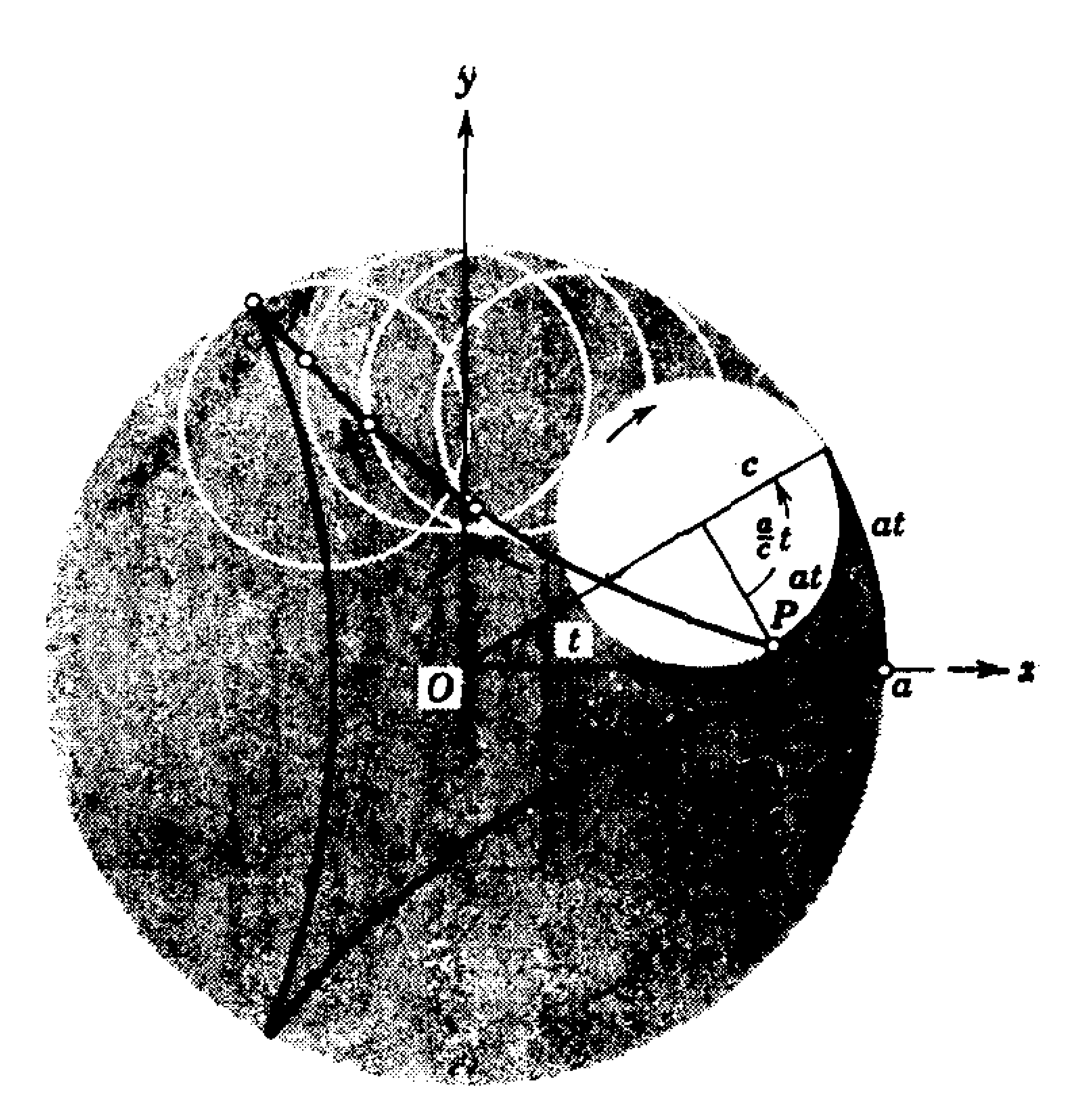


图 4.5　内摆线

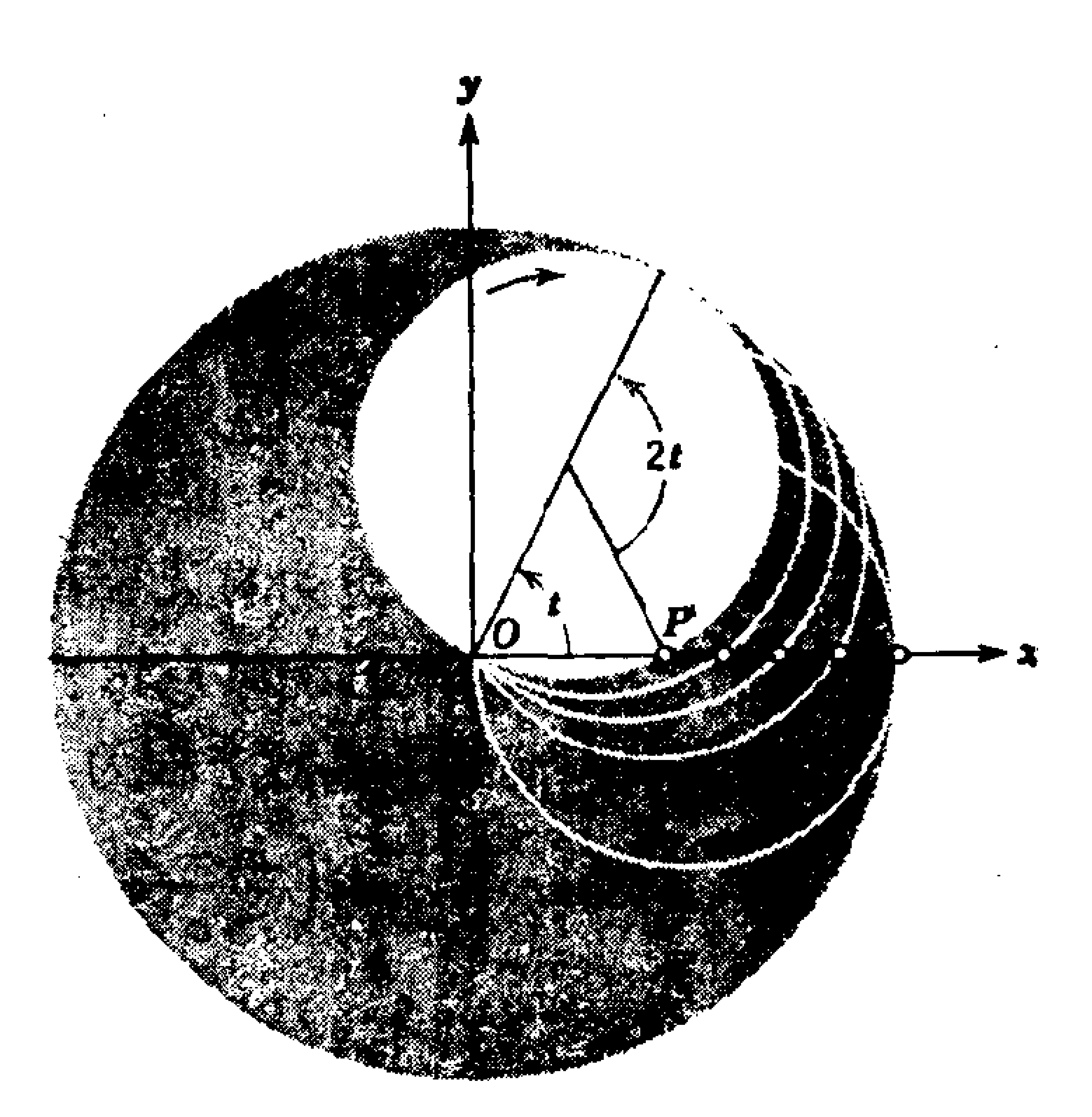


图 4.6　在 2 倍半径的圆内滚动的圆的边缘上一点 P 描绘了一条直线段

内摆线蜕化成为固定圆的直径，它不断地往复描绘．这个例子有趣的性质是，它提供了只用圆周运动画直线问题的机械方法（图 4.6）．

如果固定圆的半径为动圆半径的3倍，那么 $c=\frac{a}{3}$，因而

$$x(t)=\frac{2}{3}a\cos t+\frac{1}{3}a\cos(2t),$$

$$y(t)=\frac{2}{3}a\sin t+\frac{1}{3}a\sin(2t).$$

经过初等计算，我们有

$$x^2+y^2=\frac{5}{9}a^2+\frac{4}{9}a^2\cos(3t),$$

因此，内摆线与固定的圆刚好交于三个点，曲线如图 4.5 所示。

次摆线

如果我们考虑当一个圆沿一条直线或沿另一个圆的外边或里边滚动时，附着在该圆上的一点 P（不一定在圆的边缘上）的运动，就得到更一般的曲线，叫做次摆线（长短辐圆外旋轮线，长短辐圆内旋轮线）（见图 4.7）。当一个圆的中心本身沿着一条直线或圆匀速地移动时，在这个圆上一个匀速运动着的点的路径产生同样类型的曲线。这些曲线在行星视运动的托勒密描述中起着中心的作用。

摆线的某些重要性质将在本章后面讨论（第 451 页）。

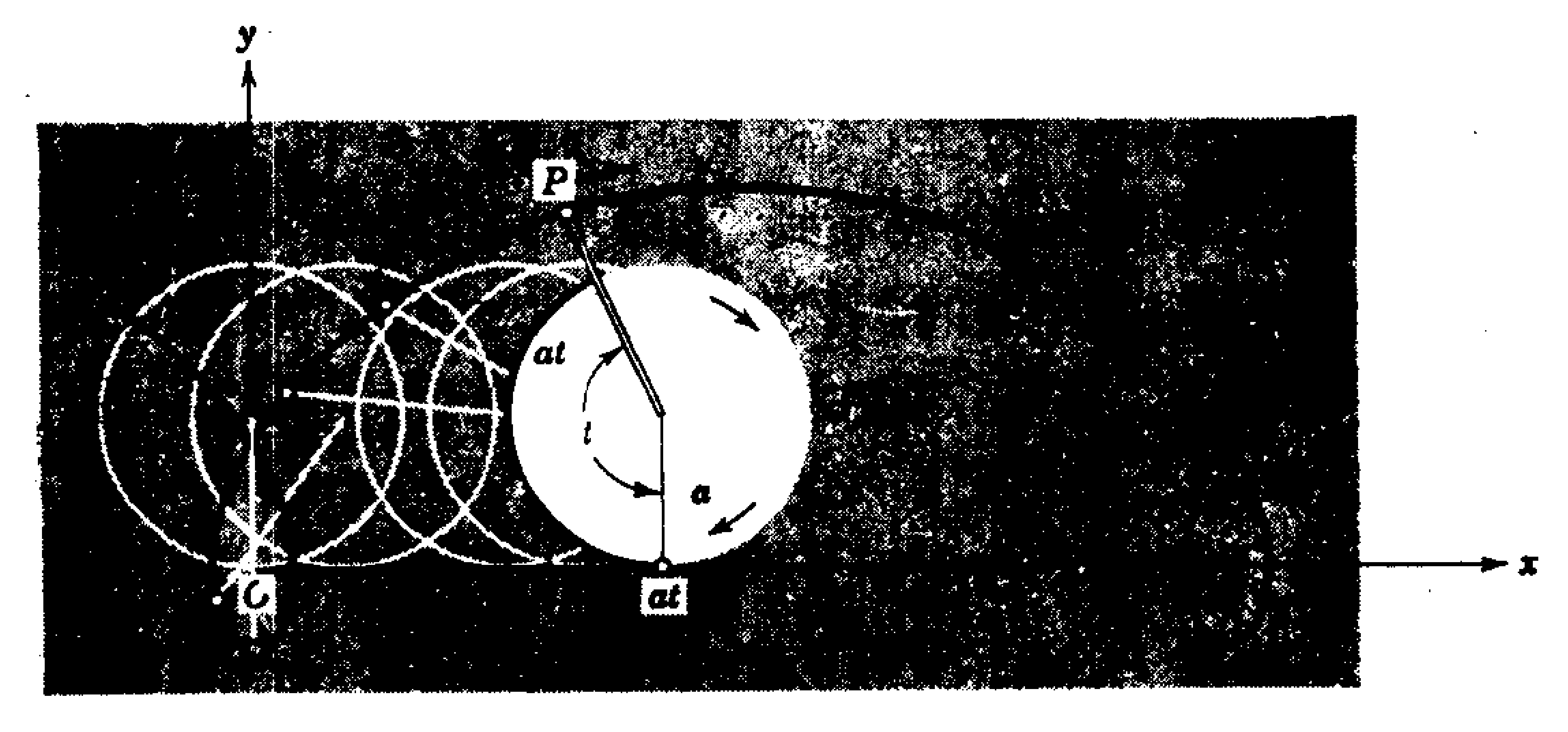


图 4.7　次摆线

d. 曲线的分类．定向

定义

曲线最明显的特性是曲线所具有的分段（即分枝）的数目

和回路的数目。双曲线是由两个不相交的分枝组成的曲线的例子；另一个例子是曲线 $y^2=(4-x^2)(x^2-1)$，它是由两个分开的卵形线组成的。我们主要讨论由一段组成的曲线，即连通的(connected)曲线。连通的曲线本身可以相交，像次摆线(图 4.7)或双纽线(109 页图 S1.3)。

一个连通曲线如果本身不相交就称为是简单的。在简单曲线中间，我们还可以区分闭的曲线，如圆和椭圆，和不闭的曲线，如抛物线或直线段。这里，我们并不企图给出严格的曲线分类，也不想给出完全的分类。而仅仅指出与参数表示有关的曲线的某些拓扑性质。

简单弧

曲线 C 的两个连续函数 $x=\phi(t)$，$y=\psi(t)$ 的参数表示定义了 t 轴或 t 轴的一部分到 C 的一个映射。如果函数 $\phi(t)$，$\psi(t)$ 的定义域是 t 轴上的闭区间 $[a,b]$，而且这个区间上不同的 t 值对应曲线 C 上不同的点 P，那么我们称 C 为简单弧。抛物线弧 $x=t$，$y=t^2$，$0\leqslant t\leqslant 1$ 是简单弧的一个例子。

可以用多种方法参数地表示同一个弧 C(即平面上同样的点)。任一单调连续函数 $\tau=\chi(t)$，$a\leqslant t\leqslant b$，定义了一个参数 τ，使得在适当的闭区间 $[\alpha,\beta]$ 上，x 和 y 是 τ 的连续函数，不同的 τ 值对应不同的 P。事实上，容易看到，连续单调的替换 $\tau=\chi(t)$ 提供了简单弧最一般的连续的参数表示，使得对不同的参数值，确定了弧的不同的点。(见第57页注，关于一对一的连续映射的说明。)

对于简单弧 C 的一个特殊的参数表示 $x=x(t)$，$y=y(t)$，C 有一个确定的指向，相应于 t 增加的方向。给定任意两个不同的点 P_0，P_1，如果 P_1 属于参数 t 较大的值，我们就说 P_1 在 P_0 的后面。如果我们用连续的增函数 $\tau=\chi(t)$ 引入新参数 τ，那么，对应于 τ，一对点的次序是同样的；参数 τ 定义了 C 的同一指向。如果 χ 是递减的，则指向就倒转了。

弧的方向或定向

有向的或定向的简单弧是这样一种弧，它有选择好的确定的

指向(例如，指向对应于特殊选择的参数 t 增加的方向)；这个指向就叫做弧的正指向．如果我们知道弧的两个端点中哪一个在另一个的后面，那么正指向就完全确定了．我们把在后面的端点称为弧的终点，另一个端点就称为起点．给定有向弧的任意参数表示 $x=x(\tau)$，$y=y(\tau)$，其中 $a\leqslant\tau\leqslant b$．如果参数值 $\tau=a$ 对应起点，$\tau=b$ 对应终点，那么正指向就是 τ 增加的方向；否则 τ 增加的方向是弧的负指向(图 4.8)．

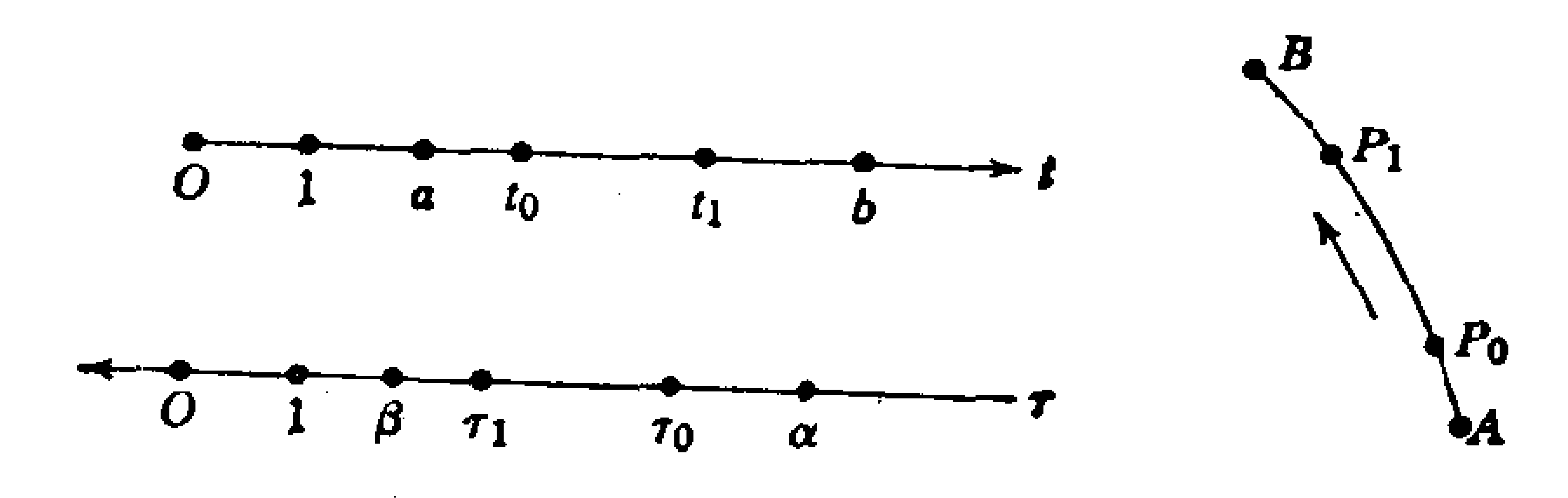


图 4.8　指向和参数表示

在一个简单弧 C 上任意两个不同的点 P_0，P_1 定义一端点为 P_0，P_1 的子弧，它由参数值介于 P_0 和 P_1 之间的那些点组成．如果 C 是有向弧，并且对于 C 的正指向 P_1 在 P_0 的后面，那么我们得到起点为 P_0，终点为 P_1 的有向子弧．有向简单弧 C 上的有限个分点把弧 C 分割成有向子弧的一个序列，一个子弧的起点是前面一个子弧的终点．

如果只限于简单弧并且硬要不同的参数值 t 一定属于曲线的不同点，那常常是不实际的．比如说，如果方程 $x=x(t)$，$y=y(t)$ 给出运动的质点 P 在时刻 t 的位置，那么没有任何理由认为质点一定不能停留一会儿，或者质点的路径一定不许和自己相交而使质点在较晚的时候再回到同样的位置．

曲线 $x=t^2-1$，$y=t^3-t$ 是一个例子(也可用三次方程 $y^2-x^2(1+x)=0$ 完全地描述它)．当 t 由 $-\infty$ 变成 $+\infty$ 时，在 $t=-1$ 和 $t=+1$ 曲线两次通过原点(图 4.9)．容易证明曲线所有的其他的点都有唯一的 t 值．在几何上区间 $-1<t<+1$ 对应于曲线的一个回路．至少，如果我们以某种方式想像出对应于 $t=-1$ 和 $t=+1$ 的点是不同的，一个点在另一个点

的"上面"，那么 t 增加的指向还是定义了曲线上点的次序．整个这个有向三次曲线可以分解为有向的简单弧，例如，分成对应于 $n \leqslant t \leqslant n+1$ 的弧，其中 n 取所有整数．

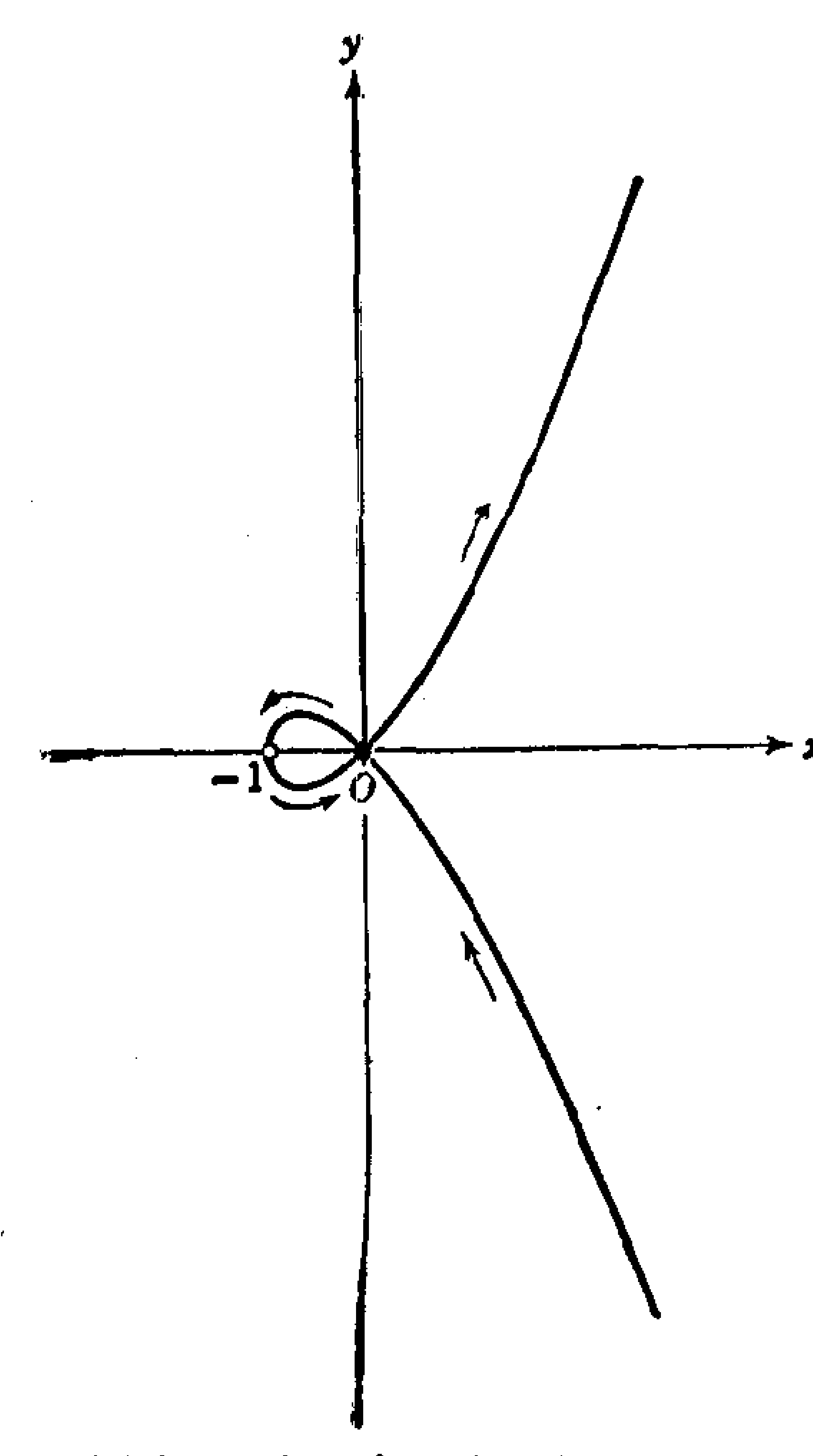


图 4.9　有一个回路的曲线：$x = t^2 - 1$，$y = t^3 - t$，指向为 t 增加方向

闭曲线

不同的 t 值对应曲线上相同的点的参数表示的一个标准例子可由公式

$$x = a\cos t, \quad y = a\sin t$$

给出，它描述了一圆上一点的匀速运动，t 为时间．当 t 从 $-\infty$ 变到 $+\infty$ 时，点 $P=(x, y)$ 以反时针方向无穷多次地画圆．如把 t 限制在长为 2π 的任意半开区间：$\alpha \leqslant t < \alpha + 2\pi$ 上，我们就刚好一次画出圆的点．区间的端点 α 和 $\alpha + 2\pi$ 对应着圆上同一个点．这里，参数区间的端点对于曲线没有特殊的几何意义．

一般地，一对连续函数 $x = \phi(t)$，$y = \psi(t)$ 定义于闭区间 $a \leqslant t \leqslant b$ 上，如果 $\phi(a) = \phi(b)$，$\psi(a) = \psi(b)$，那么它表示一条闭曲线．如果当 $a \leqslant t < b$，不同的 t 值对应不同的点 (x, y)，那么这闭曲线是简单的．

对应于 $t = a$ 和 $t = b$ 的点可以是曲线上任意一点．它正是我们"切断"曲线并使之与轴上一区间对应的那个断点．

用周期函数表示的闭曲线

正像在圆的例子中那样，使用周期 $p = b - a$ 的周期函数 $\phi(t)$ 和 $\psi(t)$，我们可以避免区分任何特殊的分割，对于周期函数，我们在这里一般地讲一讲是值得的，在第八章我们将进一步展开，

如果函数 $f(t)$ 对所有的 t 有定义，并且满足方程 $f(t)=f(t+p)$，那么 $f(t)$ 就称为周期为 p 的周期函数．例如三角函数 $\sin t$ 和 $\cos t$ 是一个周期为 2π 的周期函数．（任意倍数 $2n\pi$ 也是周期，n 为整数.）从几何解释看，如果 $f(t)$ 的图形向右移动 p 个单位而再次得到同样的图形，那么 $f(t)$ 就有周期 p.

因为 $f(t)$ 不断重复，所以如果仅知道长为 $p=b-a$ 的单个区间 $a\leqslant t<b$ 的 $f(t)$，那么对所有的 t 就决定了周期为 p 的函数 $f(t)$（图 4.10）. 事实上，对每一个 t，存在一个在区间 $a\leqslant t'<b$ 上的 t' 值，使得 $t-t'=np$，其中 n 为整数（我们只须把 n 看成不超过 $\frac{(t-a)}{p}$ 的最大整数即可）．那么就有 $f(t)=f(t')$.

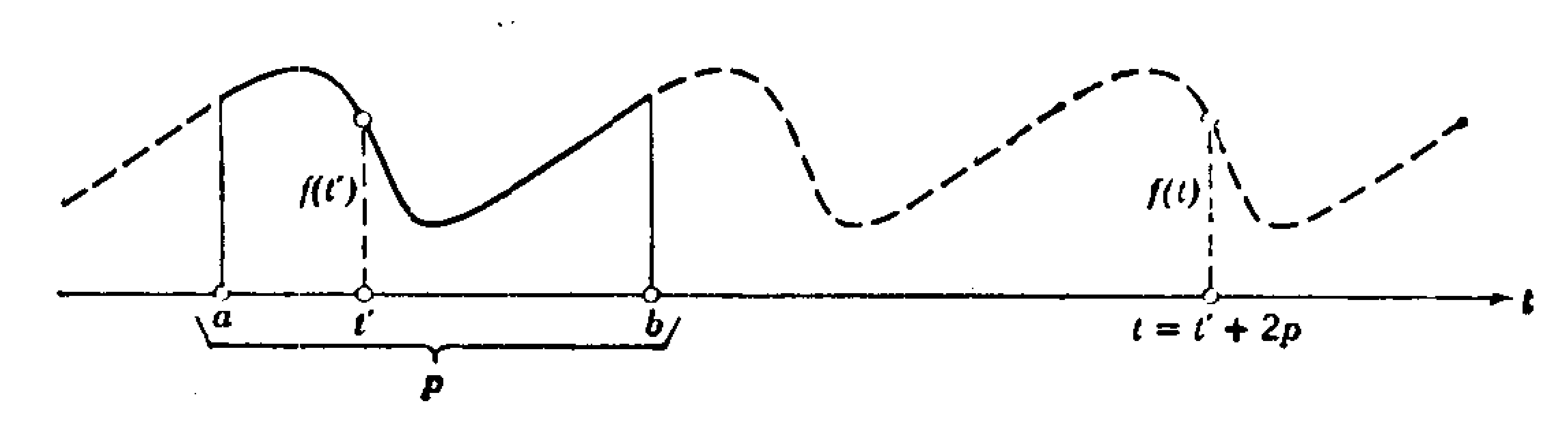


图 4.10　周期函数 $f(t)$ 的图形

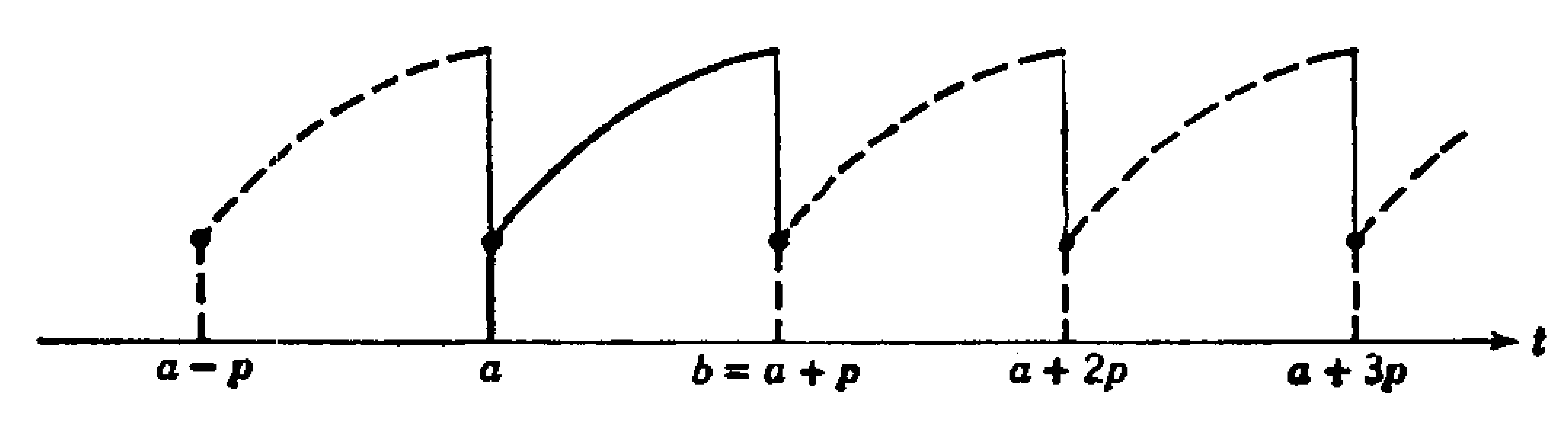


图4.11　函数 $f(t)$ 从区间 $a\leqslant t<b$ 周期地延拓

实际上，我们可以从定义在半开区间 $a\leqslant t<b$ 上的任意连续函数 $f(t)$ 着手；被延拓的函数对除 $t=a+np$ 值（n 为整数）之外的所有的 t 显然是连续的（图 4.11）.

例如，周期地延拓一个用 $f(t)=t\ (0\leqslant t<1)$ 定义的函数 $f(t)$，得到周期 $p=1$ 的函数称为"t 的小数部分"，这个函数在 t 为整数的点是不连续的（图 4.12a）．一般地，在 $t=a+np$ 时，周期延拓了的函数 f 将取值 $f(a)$；这也是从右边趋于该点时 f 的极

限，而从左边趋于该点时 f 的极限和点 b 的值相同. 现在，我们最感兴趣的是，考虑这样一个函数，它在闭区间 $a \leqslant t \leqslant b$ 上有定义并且连续，而且在端点有相同的值 $f(a) = f(b)$. 周期地延拓这样的函数总可以得到周期为 $p = b - a$ 而且对所有的 t 都连续的函数 $f(t)$ (图 4.12b).

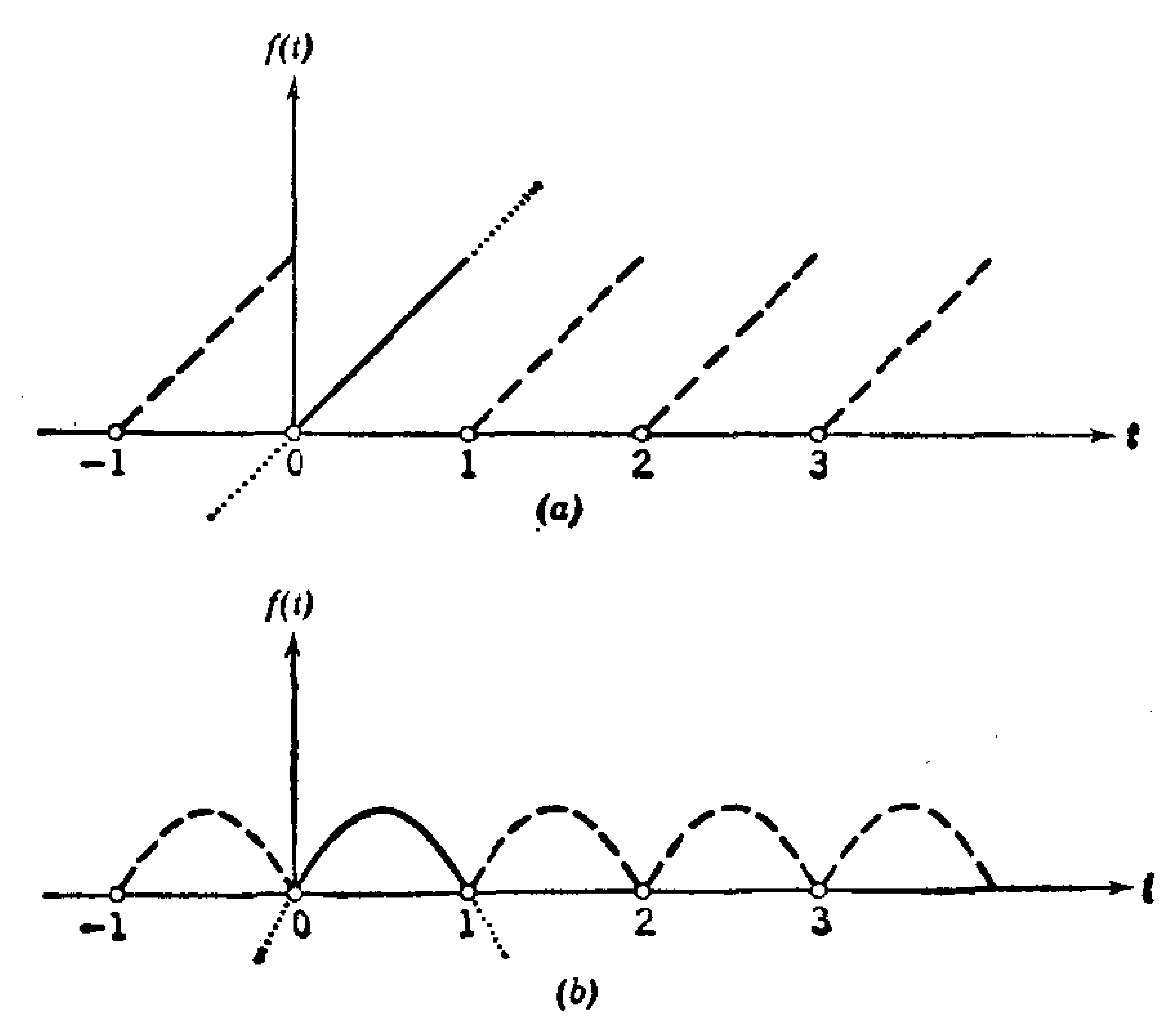


图 4.12　从区间 $0 \leqslant t < 1$ 周期地延拓函数 $f(t)$. 这里，(a) $f(t) = t$, (b) $f(t) = 2t - 2t^2$

对于表示闭曲线 C，连续的周期函数是理想的. 设 C 是用参数方程 $x = \phi(t)$, $y = \psi(t)$ 给定的，在区间 $a \leqslant t < b$ 上，ϕ, ψ 连续并在两个端点上有相同的值. 我们可以把这些函数的定义这样延拓到所有的 t 值，使 ϕ 和 ψ 的周期为 $p = b - a$，且对所有的 t 连续. 对任意的 t，被延拓了的参数表示只产生 C 的点，因为我们有 $t = t' + np$，n 为整数，$a \leqslant t' \leqslant b$. 这时，对应于 t 的点和对应于 t' 的点是位于 C 上的同一个点. 当 t 从 $-\infty$ 变到 $+\infty$ 时，点 (x, y) 无限次地通过曲线 C，就像在圆 $x = a\cos t$, $y = a\sin t$ 的情况一样. 这里参数值 $t = a$ 的特性已经消失. 对任意的 α，当 t 从 α 变到 $\alpha + p$ 时，用 $x = \phi(t)$, $y = \psi(t)$ 就已表示出整个的曲线.

如果闭曲线 C 的一部分对应于参数值 t 的区间 $\alpha \leqslant t \leqslant \beta$，在此区间上，不同的 t 值得到不同的点 (x, y)，那么这一部分就构成简单弧. 如果在同一区间 $\alpha \leqslant t < \alpha + p$ 上的不同的 t 总得到 C 上不同的点，那么整个闭曲线是简单曲线. 因此，长度小于 p 的任意闭参数区间给出简单弧.

简单弧组成的闭曲线. 点的次序

我们考虑可以分解为若干个简单弧的闭曲线. 如果整个的闭曲线是简单的，那么它可以分成两个简单弧 $t_0 \leqslant t \leqslant t_1$ 和 $t_1 \leqslant t \leqslant t_0 + p$，这两个简单弧只有它们的端点 P_0, P_1 是公共的. t 增加的方向确定了 C 的每一简单弧的正向，从而确定了 C 的正指向或**正定向**. 简单闭曲线 C 上任意两个不同的点 P_0, P_1 把 C 分成两个简单弧. 在 t 增加的指向上，两弧中恰有一个以 P_0 为起点，P_1 为终点. 我们称为 P_0P_1；对另一个弧，正好相反.

定向和次序

C 的正定向也可由有序的三点 $P_0P_1P_2$ 来描述，条件是我们要

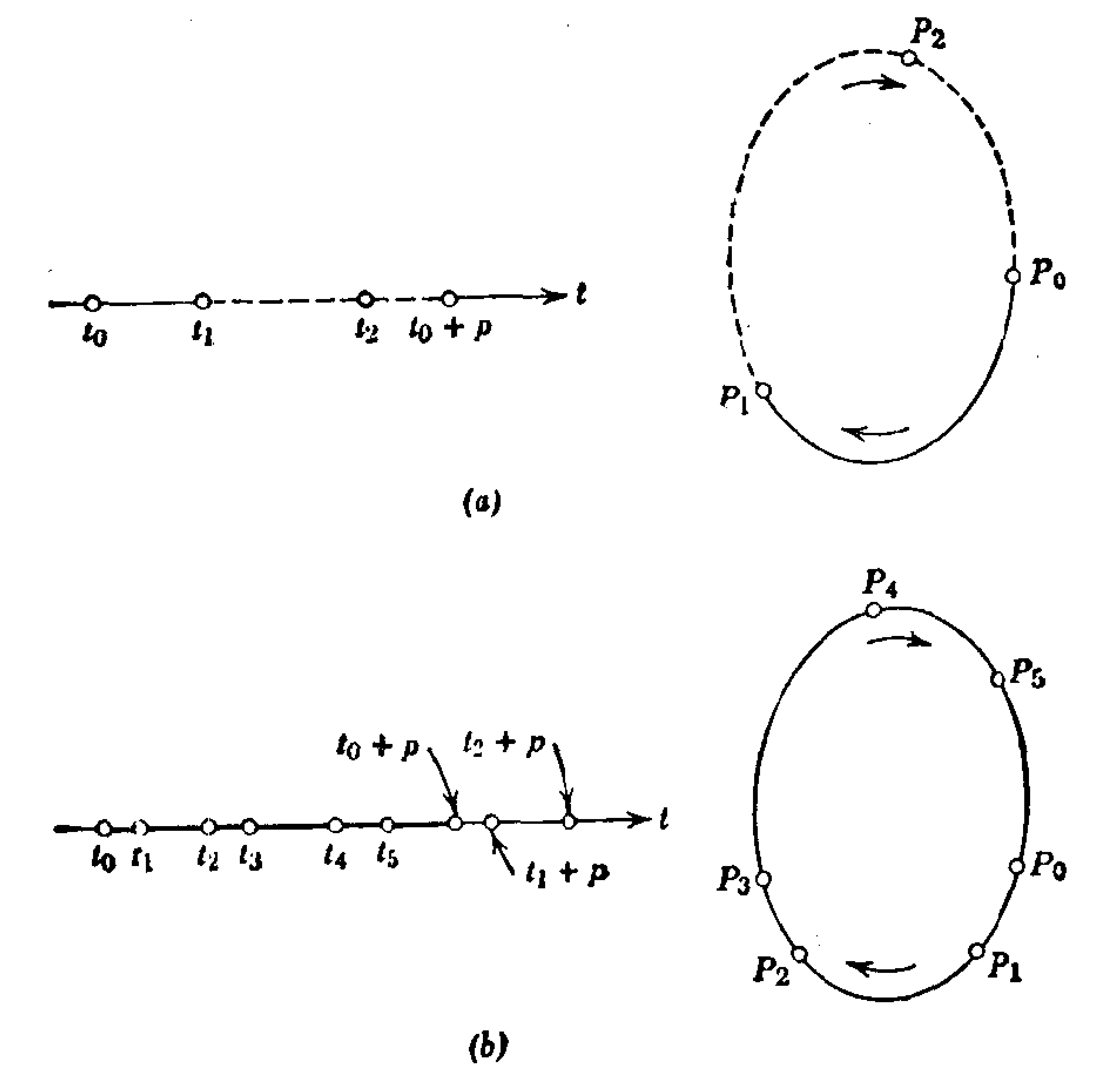


图 4.13　闭曲线按 t 增加的指向定向

指定 P_2 不在起点为 P_0 终点为 P_1 的有向简单弧上．由 $P_0P_1P_2$ 循环排列而得到的三点 $P_1P_2P_0$，$P_2P_0P_1$ 描绘了相同的定向(图 4.13a)．

*很一般地，在有向的简单闭曲线 C 上任意 n 个不同的点总是按着确定到循环排列[1]的一个次序 $P_1P_2\cdots P_n$ 分布的，并且把 C 分成有向的简单弧 P_1P_2，$\cdots$，$P_{n-1}P_n$，P_nP_1．我们总可以选择诸点 $P_1,P_2,\cdots,P_n$ 的参数值 $t_1,t_2,\cdots,t_n$ 使得 t_i 构成一个单调的递增序列，并且所有的 t_i 都包含在长度等于周期 p 的同一参数区间内(图 4.13b)．

曲线和角的定向

正如在第一章所强调的那样，我们不得不使用正号或负号来建立几何对象和以数表示的分析概念之间的令人满意的关系．最简单的例子是诸如数轴的有向线．开始时，我们把直线的哪一个方向定义为正是任意的．对于直线的任意特殊参数表示 $x=at+b$，$y=ct+d$，就能有一个相应于 t 增加的正指向与它对应．以这样方式定向的直线指向一个确定的方向．两条平行的有向直线的方向或者相同，或者相反．也可以用从一点 P_0 出发的射线来定义方向，射线即半直线，是由直线上沿正方向跟在给定点 P_0 后面的那些点组成的．

在平面上任何方向可以用从原点出发的一个射线表示，也可

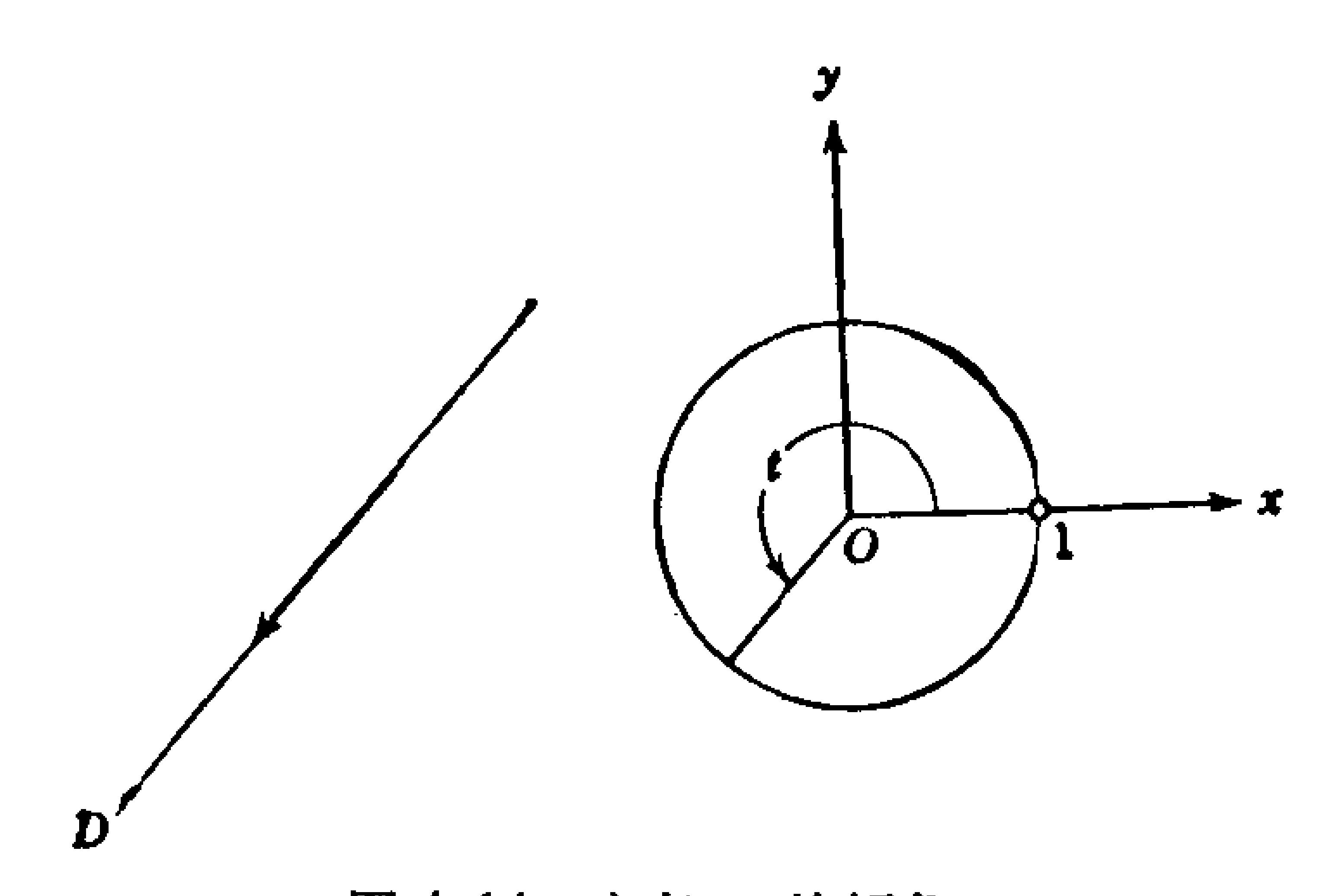


图 4.14　方向 D 的倾角 t

1) 即 $P_2P_3\cdots P_nP_1$，$P_3P_4\cdots P_nP_1P_2$，$\cdots$，$P_nP_1\cdots P_{n-1}$ 给出相同的定向．

以用该射线上的原点为中心半径为 1 的圆上的点 P 表示. 如果我们以参数方程 $x = \cos t$, $y = \sin t$ 表示这个单位圆,那么我们把每一个方向对应于某些 t 值,它们相互只差 2π 的倍数. 我们把它们称为该方向的倾角,或称为该方向与正 x 轴的夹角. 在 $0 \leqslant t < 2\pi$ 区间内总是刚好存在一个倾角 t (图 4.14).

两个方向之间的角就是它们的倾角之差. 更确切地,因为我们取两个方向的次序是有关系的,所以我们说倾角为 t' 的方向与倾角为 t'' 的方向形成夹角 $\alpha = t' - t''$ (图 4.15). 因为 t 和 t'

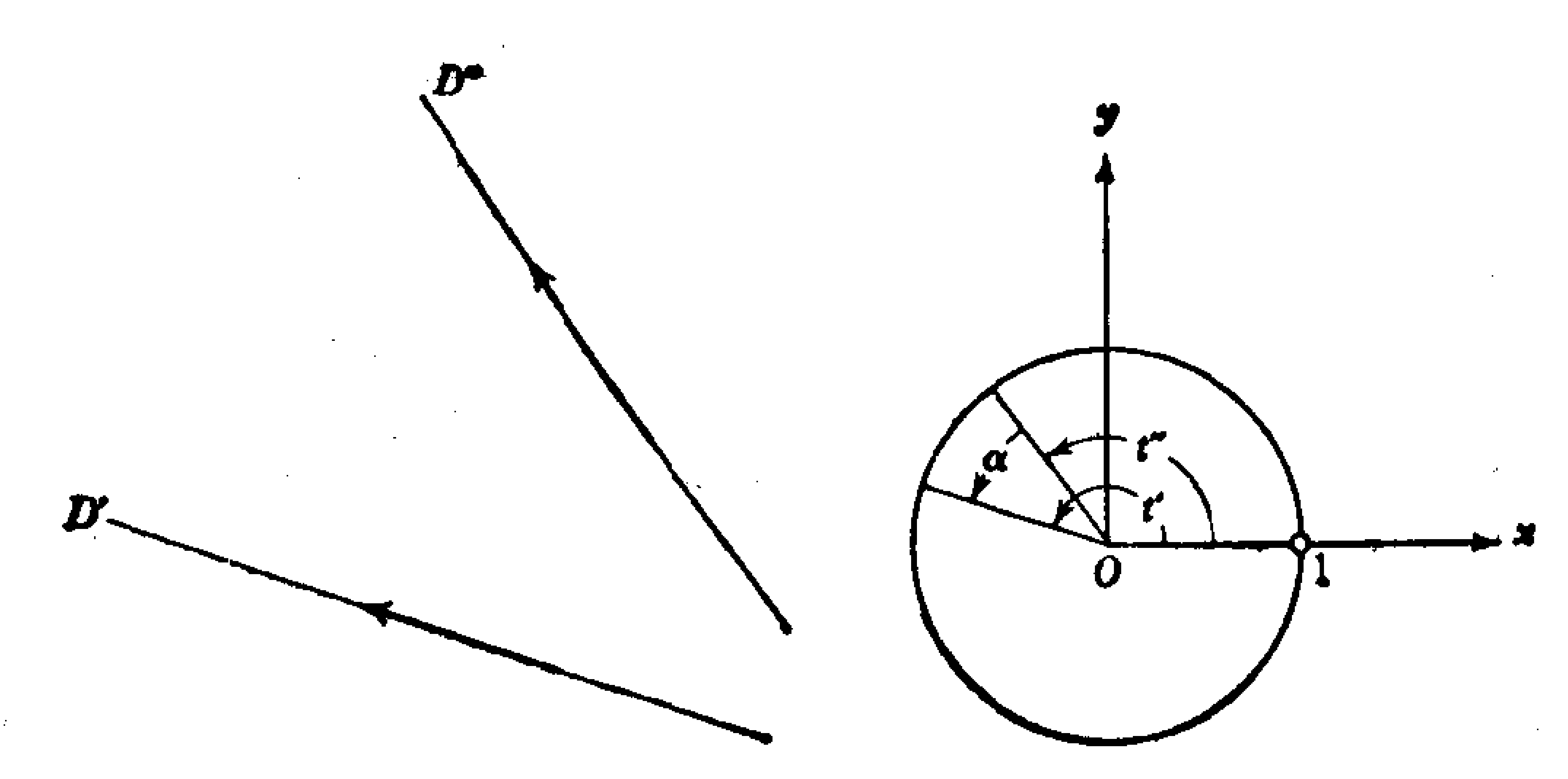


图 4.15 方向 D' 和方向 D'' 的夹角 α.

可以改变 2π 的整数倍,所以一个方向和另一个方向的夹角也可以改变 2π 的整数倍.

旋转的方向

我们也可以说倾角为 t'' 的方向经旋转 α 角而变到方向 t'. 这里,旋转的直观概念是连续运动的概念. 按此概念,倾角为 t'' 的方向变到倾角为 t' 的方向是经过了 t'' 和 t' 之间所有可能的倾角为 t 的方向. 如果 $\alpha = t' - t''$ 是正的,我们把旋转称为正的或逆时针的,否则称为负的或顺时针的. 当然,由一个给定的方向变到另一个给定的方向,可能有许多不同的顺时针和逆时针的旋转,除非我们令旋转角 α 满足 $-\pi < \alpha \leqslant \pi$.

最后,旋转的正指向与我们已经选定的圆的特殊参数表示 $x = \cos t$, $y = \sin t$ 有关. 如果 x 轴照例指向右方,y 轴指向上

方，那么旋转的正指向与平常的钟表针的方向[1]相反.

曲线的正侧和负侧

曲线把平面在曲线的一点 P 附近的点分成两类. 至少在局部

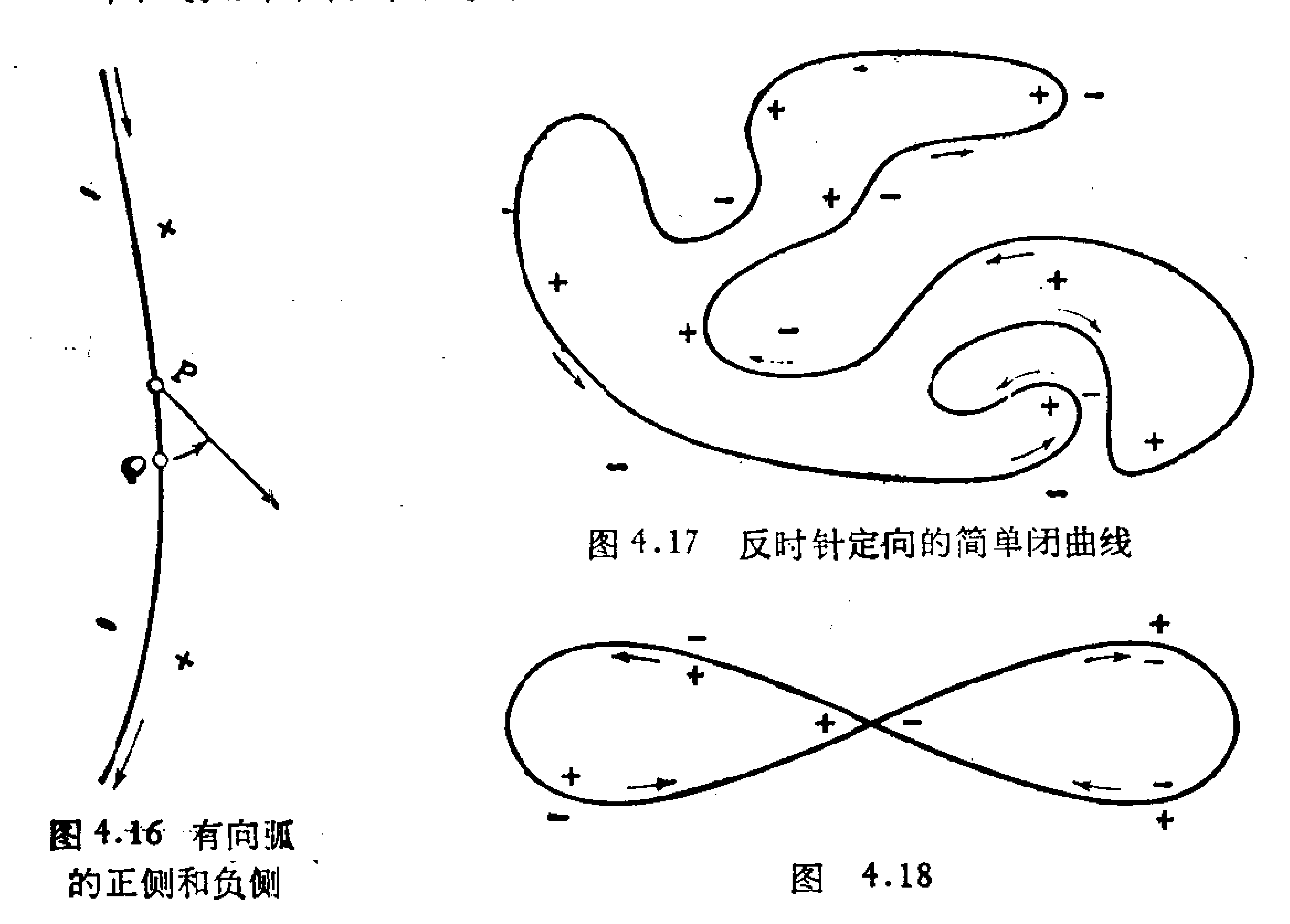


图 4.16 有向弧的正侧和负侧

图 4.17 反时针定向的简单闭曲线

图 4.18

上，我们可以区别曲线的两"侧". 如果曲线 C 是有向的，那么我们可以这样定义**正侧**("左侧")和**负侧**("右侧")[2]: 考虑从 P 发出的一条射线. 如果沿曲线给定的方向，在 P 点的后面存在任意接近 P 的点 Q，使得从 P 到 Q 的线按逆时针的方向转到给定的射线所通过的角是在 0 和 π 之间(图 4.16)，我们就说这个射线指向曲线的正侧. 在射线上接近 P 的那些点就说是在曲线的正侧. 相反的情况，就说射线指向 C 的负侧，在射线上的点就说是在曲线的负侧. 如果曲线 C 是简单的闭曲线，那么它把平面上所有的点分成在 C **内部的**和在 C **外部的**两类[3]. 如果曲线的内部在正(左)侧，那么我们就说 C **有逆时针的**定向(图 4.17).

然而，如果闭曲线 C 是由几个回路组成，那么并不是总能描绘

1) 这个指向也可由北半球日规盘影子的运动表示.

2) 术语"左侧"和"右侧"相应于通常河流按它流动的方向定向的"左岸"和"右岸".

3) 这些概念以及用简单的连续闭曲线把平面分成两部分，在拓扑上要严格地分析，这里只能凭直观接受.

C 使得所有被包围的区域都在 C 的正侧（见图 4.18）。

e. 导数．切线和法线的参数表示

方向和速率

对于一条由时间参量 t 的参量表示所给定的曲线 C

$$x = x(t) = \varphi(t), \quad y = y(t) = \psi(t),$$

按照牛顿的做法，我们用点表示导数：

$$\dot{x} = \frac{d\varphi}{dt} = \dot{\varphi}, \quad \dot{y} = \frac{d\psi}{dt} = \dot{\psi}.$$

导数 $\dot{x}$，$\dot{y}$ 通常想像为点 P 沿 C 运动的坐标的“速度分量”或“速率”。

当 $\dot{x} \neq 0$ 时，总能够用一个方程 $y=f(x)$ 表示 C 的相应的部分，办法是首先由第一个方程算出 t 为 x 的函数，然后把得到的 t 的表示式代入到第二个方程．由微分的锁链法则和反函数微分法则（见第 231 页），我们得到曲线切线的斜率

$$\frac{dy}{dx} = \frac{dy}{dt}\frac{dt}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{\dot{y}}{\dot{x}}.$$

如果 $\dot{y} \neq 0$，等价的公式 $\frac{dx}{dy} = \frac{\dot{x}}{\dot{y}}$ 成立。

除非讲到相反的情况，我们总是假定 $\dot{x}$ 和 $\dot{y}$ 不同时为零．或简单地，我们假定

$$\dot{x}^2 + \dot{y}^2 \neq 0.$$

那么切线总是存在的；[1]如果 $\dot{y} = 0$，切线是水平的，如果 $\dot{x} = 0$，它则是垂直的。

例如，对于摆线（见第 353 页式(1)），我们有

1）我们注意到，虽然条件 $\dot{x}^2 + \dot{y}^2 \neq 0$ 对于保证非参数表示是充分的，但不是必要的．例如我们可以用参数方程 $x = t^3$，$y = t^6$ 定义曲线 $y = x^2$．在 t 轴的原点，$\dot{x}^2 + \dot{y}^2$ 为正的条件不满足，但是曲线仍有定义和定义得很好的非参数表示．

$$\dot{x} = a(1 - \cos t) = 2a\sin^2\frac{t}{2},$$

$$\dot{y} = a\sin t = 2a\sin\frac{t}{2}\cos\frac{t}{2},$$

$$\frac{dy}{dx} = \cot\frac{t}{2}.$$

这些公式表明除去 $t = 0$，$\pm 2\pi, \pm 4\pi, \cdots$ 之外，$\dot{x}^2 + \dot{y}^2 \neq 0$。而且摆线在那些例外的点上有垂直切线的尖点（即曲线方向倒转的点）. 在尖点摆线还与 x 轴相交，即 $y = 0$. 在趋于这些点时，导数 $y' = \frac{\dot{y}}{\dot{x}} = \cot\frac{t}{2}$ 变为无穷大。

切线，法线和方向余弦

曲线在点 x，y 的切线方程是

$$\eta - y = \frac{dy}{dx}(\xi - x),$$

其中 ξ，η 是对应于切线上任意一点的"流动 (running)"坐标，而 x，y 和 $\frac{dy}{dx}$ 取由切点决定的固定值. 以 $\frac{\dot{y}}{\dot{x}}$ 代 $\frac{dy}{dx}$，我们可以把切线方程写成

$$(\xi - x)\dot{y} - (\eta - y)\dot{x} = 0. \tag{5}$$

在假定 $\dot{y} \neq 0$ 的条件下，我们只须把 x 表示为 y 的函数就得到完全一样的方程. 在那些例外的点上，$\dot{x}$ 和 $\dot{y}$ 对同一 t 值都等于零，这个方程变得没有意义，因为它对于所有的 ξ，η 都满足.

过曲线上一点并与这个点的切线垂直的直线叫做曲线的法线. 法线的斜率是 $-\frac{dx}{dy}$. 因此，法线方程为

$$(\xi - x)\dot{x} + (\eta - y)\dot{y} = 0. \tag{6}$$

如果 C 的一点对应 t 的几个值，那么一般地说，对于通过这个点的曲线的每一个分枝，或者说对于 t 的每个值，就有不同的切线. 例如，曲线 $x = t^2 - 1$，$y = t^3 - t$ (第 360 页图 4.9)有 $t = -1$ 和 $t = +1$ 两个值通过原点. 对 $t = -1$，我们得到切线方程

$\xi+\eta=0$，而 $t=+1$ 的切线方程为 $\xi-\eta=0$.

由导数的定义，我们有

$$\frac{dy}{dx}=\frac{\dot{y}}{\dot{x}}=\tan\alpha,$$

其中 α 为切线与 x 轴的夹角．这意味着把 x 轴旋转角 α（如 $\alpha>0$，沿逆时针方向转；如 $\alpha<0$，沿顺时针方向转）会使它与切线平行．以角度 $\alpha\pm\pi$，$\alpha\pm2\pi$，$\cdots$ 旋转 x 轴也会使它平行于切线，所以角度 α 确定到仅差 π 的整数倍，而 $\tan\alpha$ 是唯一确定的．由关系式 $\dot{y}/\dot{x}=\sin\alpha/\cos\alpha$ 和 $\dot{x}^2+\dot{y}^2\neq0$，我们有

$$\cos\alpha=\pm\frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}},\qquad \sin\alpha=\pm\frac{\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}},$$

这里两式必须取相同符号．我们把 $\cos\alpha$ 和 $\sin\alpha$ 叫做切线的**方向余弦**[1]．

切线和法线的指定方向

方向余弦的两个可能的选择对应着我们可以作切线的两个方向；对应的角 α 可相差 π 的奇数倍．切线上两方向之一对应 t 增加的方向，另一个对应 t 减少的方向．假定曲线的指向是 t 增加的方向，那么，按照定义，切线的正方向或者说相应于 t 值增加的方向是与正 x 轴夹角 α 的方向，而对于角 α，$\cos\alpha$ 与 $\dot{x}$ 同号，且 $\sin\alpha$ 与 $\dot{y}$ 同号．在切线上该方向的方向余弦无疑是

$$\cos\alpha=\frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}},\quad \sin\alpha=\frac{\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}}. \tag{7}$$

如果，譬如说，$\dot{x}=\dfrac{dx}{dt}>0$，那么在切线上 t 增加的方向即是 x 增加的方向；该方向与正 x 轴的夹角的余弦是正的．类似地，把相应于 t 增加的正切线的方向以正（逆时针）指向旋转 $\dfrac{\pi}{2}$ 得到的法线方向，它的方向余弦无疑是

1）这里，我们把 $\sin\alpha$ 看成 $\cos\beta$，其中 $\beta=\dfrac{\pi}{2}-\alpha$ 为 y 轴与切线的夹角．

$$\cos\left(\alpha+\frac{\pi}{2}\right)=\frac{-\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}},\quad \sin\left(\alpha+\frac{\pi}{2}\right)=\frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}}.$$

这个方向叫做正法线方向并指向曲线的正侧（图 4.19）。

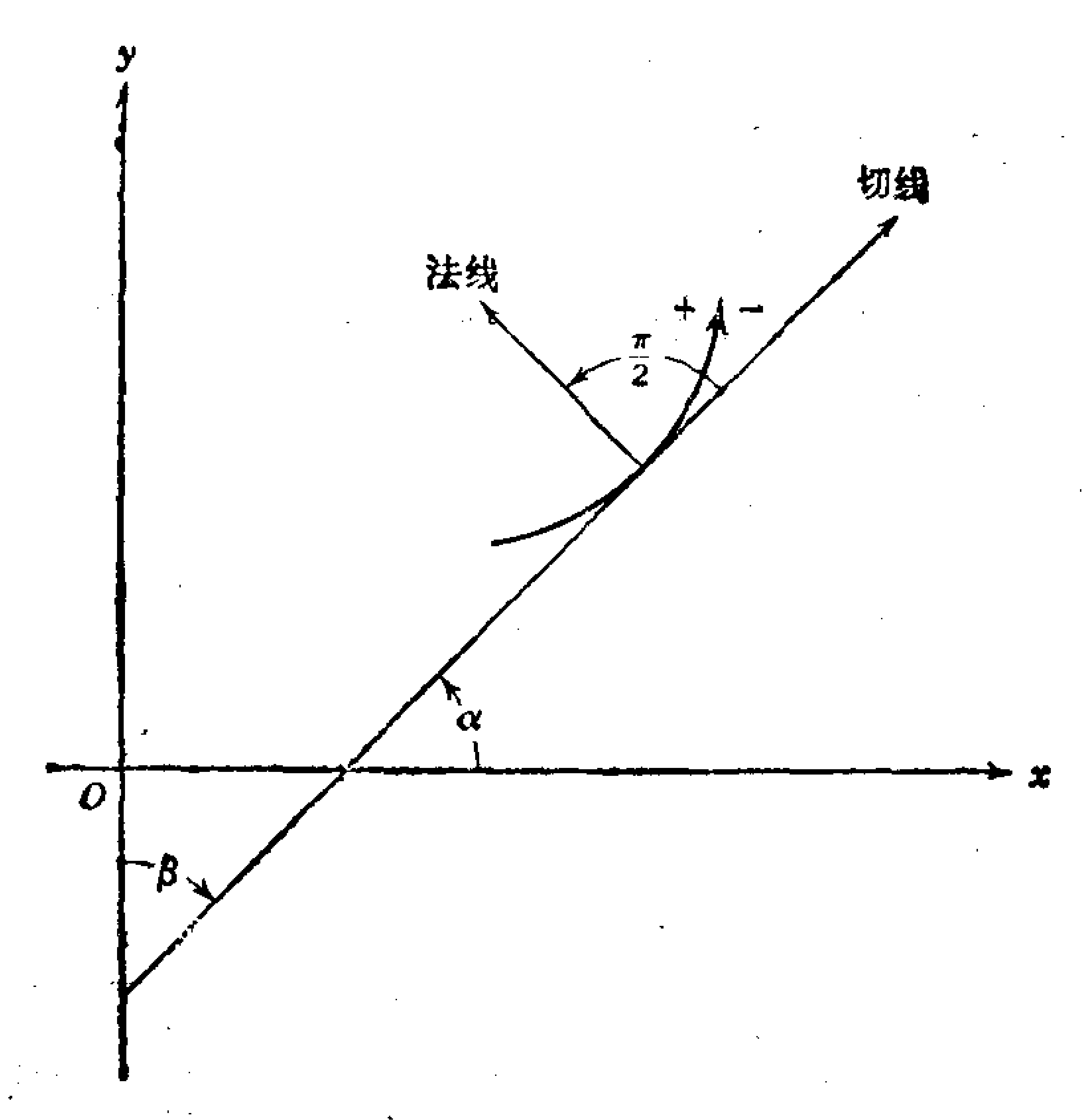


图 4.19 有向曲线的正切线和正法线

如果在曲线上我们引入新的参数 $\tau=\chi(t)$，那么当 $\frac{d\tau}{dt}>0$ 时，$\cos\alpha$ 和 $\sin\alpha$ 的值不变，当 $\frac{d\tau}{dt}<0$ 时，$\cos\alpha$ 和 $\sin\alpha$ 改变符号；就是说，如果我们改变曲线的指向，那么切线和法线的正指向要同样地改变。

临界点

如果 $\dot{x}$ 和 $\dot{y}$ 连续并且 $\dot{x}^2+\dot{y}^2>0$，那么决定切线方向的量 $\cos\alpha$ 和 $\sin\alpha$ 将随 t 连续地变化. 从而，方程为

$$(\xi-x)\sin\alpha-(\eta-y)\cos\alpha=0$$

的切线沿着曲线连续地改变，法线也同样.

如果对 t 的某个值，$\dot{x}$ 和 $\dot{y}$ 都是零，那么按我们的公式，切线的方向余弦没有定义；切线可能完全不存在或可能不唯一确定. 这样的点叫做"临界"点或"稳定"点. 我们举例说明产生临界点的各

种可能性.

一个例子是曲线 $y=|x|$，它有参数表示为 $x=t^3$，$y=|t|^3$；这个曲线在 $t=0$ 有一个角点，虽然 $\dot{x}$ 和 $\dot{y}$ 都保持连续. 在第 368 页讨论的摆线的例子中 $\dot{x}=\dot{y}=0$ 的"稳定"点对应于尖点. 另一方面，在某些情况下 $\dot{x}$ 和 $\dot{y}$ 等于 0 和曲线的性态并没有关系，而只不过是特殊的参数表示的本性，如 $x=t^3$，$y=t^3$ 表示的直线在参数值 $t=0$ 即是.

角点

由在角点上相交的几个光滑的弧所组成的曲线在参数表示中用函数 $x(t)$，$y(t)$ 来表达，它们是连续的但具有跳跃间断的导数 $\dot{x}$，$\dot{y}$. 我们用表示为

$$x=t,\quad y=0,\quad \text{对于}\quad t\leqslant 0,$$

和

$$x=t,\quad y=t,\quad \text{对于}\quad t\geqslant 0$$

的折线的简单例子来说明. 这里，对于 $t<0$，有 $\dot{x}=1$，$\dot{y}=0$；对于 $t>0$ 有 $\dot{x}=1$，$\dot{y}=1$. 在 $t=0$ 点，切线是不确定的(见图 4.20).

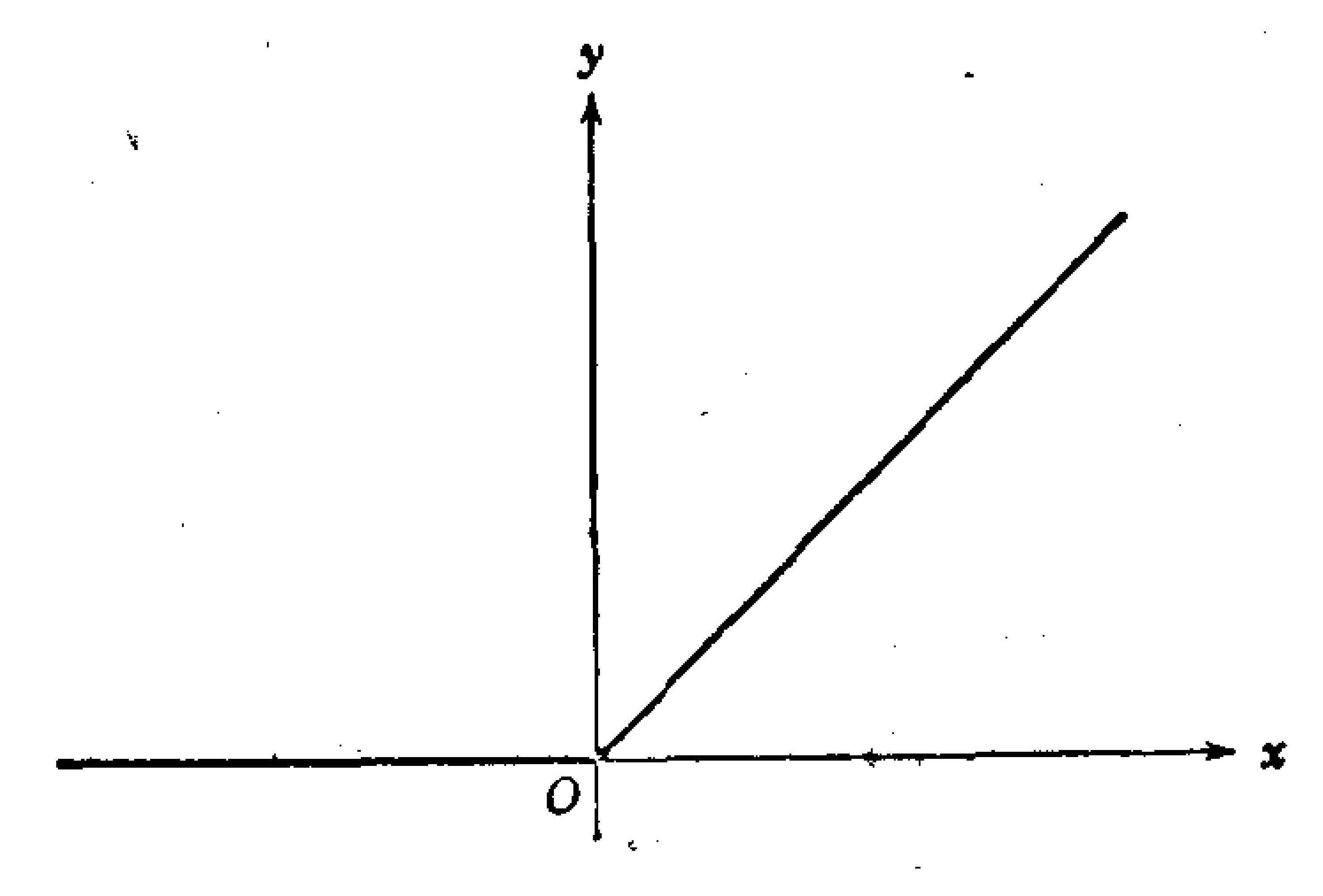


图 4.20　$x=t$，　$y=\frac{1}{2}(t+|t|)$ 的图形

f. 曲线的长度

长度当作积分

曲线有两种不同类型的几何性质或几何量. 第一种类型只依

赖于曲线在小范围内的性态，即在一个点紧接的邻域内的性态．这样的性质可以用在一点的导数来表达．第二种类型的性质，或者说在大范围内的性质则依赖于曲线或曲线一部分的整个形状，通常用积分概念解析地表达．我们开始考虑一个第二种类型的量，曲线的长度．

当然，我们对于曲线长度的含义已有直观的概念．但是，正像在圆弧的古典情况中一样，对于直观的概念必须给出严谨的数学意义．由直观作引导，我们把任意曲线的长度定义为近似多边形长度的极限，特别是内接多边形长度的极限．长度单位一经选定，多边形的长就确定了．最后的结果将是用积分表示曲线长度的公式．

我们假设曲线由方程组 $x=x(t)$，$y=y(t)$ $(\alpha\leqslant t\leqslant\beta)$ 给出．在 α 和 β 之间的区间上，我们选择中间点 t_1，t_2，$\cdots$，t_{n-1} 使得

$$\alpha = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = \beta.$$

我们把对应于这些 t_i 值的曲线上的点 $P_0, P_1, \cdots, P_n$ 用线段依次连接起来，得到内接多边形．内接多边形的周长依赖于点 t_i 或多边形顶点 P_i 的选择方法．现在我们设点 t_i 的数目以这样方式无限地增加，使最长的子区间 (t_i, t_{i+1}) 的长度同时趋于零．曲线长度就定义为这些内接多边形周长的极限，条件是这样的极限存在并且不依赖于多边形的特殊选择方法．当这个假定（可求长假定）满足的时候，我们就可以谈得上曲线的长度．

我们假定函数 $x(t)$ 和 $y(t)$ 在 $\alpha\leqslant t\leqslant\beta$ 有连续的导数 $\dot{x}(t)$ 和 $\dot{y}(t)$．相应于用点 t_i 剖分 t 区间并且 $\Delta t_i = t_{i+1} - t_i$，内接多边形有顶点 $P_i = (x(t_i),\ y(t_i))$；根据毕达哥拉斯定理（参看下页图 4.21），这个多边形全长的表达式为

$$\begin{aligned}S_n &= \sum_{i=0}^{n-1}\overline{P_iP_{i+1}}\\ &= \sum_{i=0}^{n-1}\sqrt{[x(t_{i+1})-x(t_i)]^2+[y(t_{i+1})-y(t_i)]^2}.\end{aligned}$$

用微分中值定理得

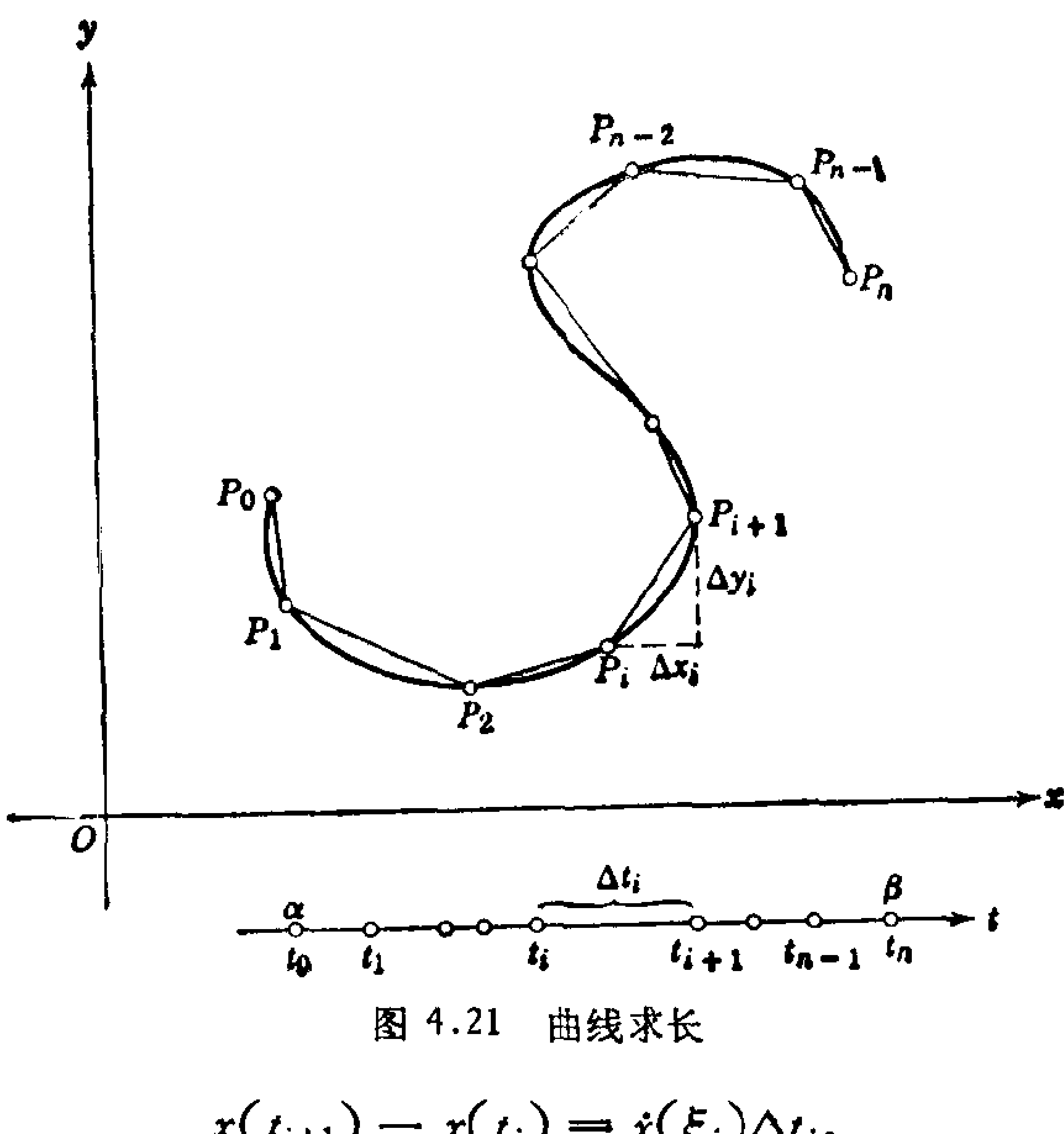


图 4.21 曲线求长

$$x(t_{i+1}) - x(t_i) = \dot{x}(\xi_i)\Delta t_i,$$
$$y(t_{i+1}) - y(t_i) = \dot{y}(\eta_i)\Delta t_i,$$

其中 ξ_i 和 η_i 是区间 $t_i < t < t_{i+1}$ 内的中间值. 因此，得到多边形的长度为

$$S_n = \sum_{i=0}^{n-1} \sqrt{\dot{x}^2(\xi_i) + \dot{y}^2(\eta_i)}\,\Delta t_i,$$

其中我们用了差 Δt_i 是正的这一事实. 如果剖分点 t_i 的数目 n 无限地增加，同时最大的值 Δt_i 趋于零，那末和 S_n 趋于积分

$$L = \int_\alpha^\beta \sqrt{\dot{x}^2 + \dot{y}^2}\,dt.$$

这个事实是第二章积分存在定理的一个直接的推论[1].

这就证明了对于连续的 $\dot{x}$, $\dot{y}$ 曲线确实有一个长度，并且这个长度由表达式

$$L = \int_\alpha^\beta \sqrt{\dot{x}^2 + \dot{y}^2}\,dt \tag{8}$$

1) 因为中间点 ξ_i 和 η_i 不一定重合，我们用更一般的近似和，在第 207 页上已表明它收敛于积分.

解析地给出。如果允许 $\dot{x}$ 和 $\dot{y}$ 在孤立点上不连续，表达式同样成立，在孤立点上曲线也许没有唯一的切线；这个积分当然必须看成为广义积分（见第三章第 321 页）。对于更一般的"可求长的"曲线，我们的积分是有意义的，但在这一卷我们将不讨论。

长度的另一个定义

我们再作一个有趣的观察：任何内接多边形 π 的周长 S 决不超过曲线的长度 L.（特别地，曲线端点的距离不能超过 L；因为连接端点的直线是连接这两点的最短的曲线.）事实上，我们可以作一个内接多边形的特殊序列，从周长为 S 的多边形 π 开始，然后逐次增加顶点而得到一序列的多边形，L 是这个周长的序列的极限，在内接多边形两个相邻的顶点中间插入一个另外的顶点不会导致周长减小，因为三角形的一个边不会超过其他两边的和. 因此，L 是由 S 开始的非减的周长序列的极限. 因之，$S \leqslant L$. 所以，如果不把 L 定义为相应于越来越细的 t 区间剖分的内接多边形序列的周长的极限，我们也可以把 L 定义为所有的内接多边形周长的最小上界. 有趣的是，不需要在形式上借助于任何取极限的过程也可以定义曲线的长度.

在参数变换下长度的不变性

根据定义，显然曲线 C 的长度 L 不能依赖于我们用于 C 的特殊的参数表示. 因此，如果我们引入一个新的参数 $\tau = \chi(t)$，其中 $\frac{d\tau}{dt} > 0$，那么，不论 t 作为参数还是 τ 作为参数，L 的积分公式必定给出相同的值. 这一点可以根据微分的锁链法则和积分的换元法来证明. 事实上

$$\begin{aligned}\sqrt{\dot{x}^2+\dot{y}^2} &= \sqrt{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2} \\ &= \sqrt{\left(\frac{dx}{d\tau}\frac{d\tau}{dt}\right)^2+\left(\frac{dy}{d\tau}\frac{d\tau}{dt}\right)^2} \\ &= \sqrt{\left(\frac{dx}{d\tau}\right)^2+\left(\frac{dy}{d\tau}\right)^2}\,\frac{d\tau}{dt};\end{aligned}$$

因之，如果 $\chi(\alpha)=a$，$\chi(\beta)=b$，我们有

$$L=\int_\alpha^\beta\sqrt{\dot{x}^2+\dot{y}^2}dt=\int_\alpha^\beta\sqrt{\left(\frac{dx}{d\tau}\right)^2+\left(\frac{dy}{d\tau}\right)^2}\frac{d\tau}{dt}\ dt$$

$$=\int_a^b\sqrt{\left(\frac{dx}{d\tau}\right)^2+\left(\frac{dy}{d\tau}\right)^2}d\tau,$$

所以基于参数 τ 的长度的表达式得到同一个值 L。如果，相反 $\frac{d\tau}{dt}<0$，类似地我们有

$$\int_\alpha^\beta\sqrt{\dot{x}^2+\dot{y}^2}dt=-\int_a^b\sqrt{\left(\frac{dx}{d\tau}\right)^2+\left(\frac{dy}{d\tau}\right)^2}d\tau$$

$$=\int_b^a\sqrt{\left(\frac{dx}{d\tau}\right)^2+\left(\frac{dy}{d\tau}\right)^2}d\tau;$$

因为 $\chi(t)$ 是减函数，而现在 $b<a$，所以右边仍然是对于参数 τ 的 C 的长度的正确的积分.

对于由函数 $y=f(x)$ $(a\leqslant x\leqslant b)$ 给定的非参数表示的曲线，我们可用 x 作为参数 t，那么 $\dot{x}=1$，$\dot{y}=\frac{dy}{dx}$. 曲线的长度由

$$L=\int_a^b\sqrt{1+\left(\frac{dy}{dx}\right)^2}dx \tag{9}$$

给出.

例

作为一个例子，我们求抛物线 $y=\frac{1}{2}x^2$ 对应于区间 $a\leqslant x\leqslant b$ 的一段的长度

$$L=\int_a^b\sqrt{1+x^2}dx.$$

这里，作代换 $x=\sinh t$（见第三章第 291 页）得到

$$\int_{\operatorname{ar\,sinh}a}^{\operatorname{ar\,sinh}b}\cosh^2 t\,dt=\frac{1}{2}\int_{\operatorname{ar\,sinh}a}^{\operatorname{ar\,sinh}b}(1+\cosh 2t)dt$$

$$=\frac{1}{2}(t+\sinh t\cosh t)\Big|_{\operatorname{ar\,sinh}a}^{\operatorname{ar\,sinh}b}$$



$$= \frac{1}{2}(\operatorname{ar}\sinh b + b\sqrt{1+b^2} - \operatorname{ar}\sin a - a\sqrt{1+a^2}).$$

在极坐标中，由方程 $r = r(\theta)(\alpha \leqslant \theta \leqslant \beta)$给定的曲线，我们有表示式 $x = r(\theta)\cos\theta$，$y = r(\theta)\sin\theta$. 取 θ 为参数，我们有

$$\dot{x} = \dot{r}\cos\theta - r\sin\theta,\quad \dot{y} = \dot{r}\sin\theta + r\cos\theta,\quad \dot{x}^2 + \dot{y}^2 = r^2 + \dot{r}^2.$$

于是对于在极坐标中曲线的长度就得到表达式

$$L = \int_\alpha^\beta \sqrt{r^2 + \left(\frac{dr}{d\theta}\right)^2}\, d\theta. \qquad (10)$$

例如，对于以原点为中心，半径为 a 的圆，我们有方程 $r =$ 常数 $= a$，$0 \leqslant \theta \leqslant 2\pi$；从而给出圆的全长是

$$L = \int_0^{2\pi} a\, d\theta = 2\pi a.$$

长度的可加性

设 C 为由 $x = x(t)$，$y = y(t)$，$\alpha \leqslant t \leqslant \beta$ 给定的曲线，其中 $\dot{x}$ 和 $\dot{y}$ 连续. 设 γ 是 α 和 β 之间的任意一个值. 由积分的一般公式，我们有

$$\int_\alpha^\beta \sqrt{\dot{x}^2 + \dot{y}^2}\, dt = \int_\alpha^\gamma \sqrt{\dot{x}^2 + \dot{y}^2}\, dt + \int_\gamma^\beta \sqrt{\dot{x}^2 + \dot{y}^2}\, dt.$$

右边的两个积分分别表示以对应于 $t = \gamma$ 的点分割 C 所成的两部分的长度. 因此，曲线的全长等于它的部分长度的和.

$\dot{x}$ 和 $\dot{y}$ 是连续的并非必要. 当 $\dot{x}$ 和 $\dot{y}$ 有有限个跳跃不连续点时，积分同样是存在的，在曲线有角点时就会出现这样的情况. 曲线的全长是各角点之间的光滑部分长度的和. 只要曲线长度的表示式作为广义积分是有意义的，那么 $\dot{x}$ 和 $\dot{y}$ 的更强的奇性也是允许的.

g. 弧长作为参数

我们已经看到，同一条曲线允许有很多不同的参数表示 $x = x(t)$，$y = y(t)$. t 的任何单调函数都可代替 t 作为参数. 然而，为了多种目的，用某种方法几何地选取“标准参数”，把曲线 C 与之

联系是有益的. 如果要在大范围描述曲线，那么横坐标 x 或极角 θ 都不适合于这种目的；而且它们依赖于坐标系的选择. 沿曲线测量长度的可能性为我们提供一个自然的用几何定义的参数，即点 P 与某一固定点 P_0 之间的曲线部分的长度，用这个参数能指出一条可求长的曲线上的各个点 P.

我们从 C 的任一参数表示 $x=x(t),\ \ y=y(t),\ \ \alpha\leqslant t\leqslant\beta$ 出发. 用点表示对 t 的微分. 我们用不定积分引入"弧长" s:

$$s=\int\sqrt{\dot{x}^2+\dot{y}^2}dt \tag{11}$$

或更确切地，把 s 表示成 t 的函数:

$$s=s(t)=c+\int_{t_0}^{t}\sqrt{\dot{x}^2(\tau)+\dot{y}^2(\tau)}d\tau, \tag{11.a}$$

其中 c 是常数，t_0 是 α 和 β 之间的一个值，并且为了与积分的上限区别，我们把积分变量记作 τ. 显然，对于参数区间中任意的值 t_1 和 t_2，差

$$s(t_2)-s(t_1)=\int_{t_1}^{t_2}\sqrt{\dot{x}^2+\dot{y}^2}d\tau \tag{12}$$

等于对应于 $t=t_1$ 和 $t=t_2$ 的点之间的曲线部分的长度，条件是 $t_1<t_2$. 对 $t_1>t_2$，差 $s(t_2)-s(t_1)$ 为该部分长度的负值. 因此，了解了任一不定积分 s，我们就能够计算曲线任何部分的长度.

弧长的正负号

如果常数 c 为零，那么我们可把 $s(t)$ 本身理解为参数为 t_0 的点 P_0 和参数为 t 的点 P 之间的曲线的弧长（或"沿曲线的距离"）；这里，当以 P_0 为起点以 P 为终点的弧的定向与 t 增加相对应时，长度算是正的[1].

s 的积分形式定义与下式等价，

$$\frac{ds}{dt}=\sqrt{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2}. \tag{12a}$$

1) 注意，变量 s 不是完全唯一的；它依赖于 P_0 和 c 的选择，也依赖于由参数 t 引进曲线的定向. 但是，任何其他的弧长都可用 s 表为(s + 常数)或($-s$ + 常数).

用微分(第 191 页)的符号表示$ds=\left(\frac{ds}{dt}\right)dt$，等等，我们可以把这个关系式写成对“长度元素” ds 的启发性的形式

$$ds=\sqrt{dx^2+dy^2}.$$

沿曲线运动的速率

如果 t 理解为时间，并且 $x(t)$，$y(t)$ 是运动着的点在时刻 t 的位置的坐标，我们有

$$\dot{s}=\frac{ds}{dt}=\lim_{h\to 0}\frac{s(t+h)-s(t)}{h}$$

为该点沿着它的路径运动的距离对于时间的变化率，即质点的速率. 对于沿曲线匀速运动的质点，$\dot{s}$ 是一个常数，并且 s 是时间 t 的线性函数.

如果满足我们通常的假定

$$\dot{x}^2+\dot{y}^2\neq 0,$$

我们有 $\frac{ds}{dt}\neq 0$，因而可以把 s 本身作为参数. 这样一来，很多公式和计算就简化了. 下列两个量

$$\frac{dx}{ds}=\frac{dx}{dt}\frac{dt}{ds}=\frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}},$$

$$\frac{dy}{ds}=\frac{dy}{dt}\frac{dt}{ds}=\frac{\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}}$$

刚好是指向 s 增加方向的切线的方向余弦（见第 369 页 (7)). 关系式

$$\left(\frac{dx}{ds}\right)^2+\left(\frac{dy}{ds}\right)^2=1, \tag{13}$$

是参数 s 作为沿曲线的弧长的特征.

h. 曲率

用方向变化率定义

下面，我们讨论一个关于曲线在一点的邻域上的局部性态的基本概念，即曲率的概念.

当我们描绘曲线的时候，曲线的倾角 α 将以一个确定的相对于所经过的单位弧长的改变率变化。α 的变化率叫做**曲线的曲率**。因此，曲率（curvature）定义为

$$\kappa = \frac{d\alpha}{ds}. \tag{14}$$

参数表达式

设曲线由函数 $x=x(t), y=y(t)$ 以参数形式给出，它们对 t 有连续的一阶和二阶导数，并且 $\dot{x}^2+\dot{y}^2 \neq 0$. 在计算方向角 α 在 P 点的改变率时，我们必须注意 α 并不唯一确定. 但是，α 的三角函数 $\tan\alpha = \frac{\dot{y}}{\dot{x}}$ $\left(\text{或 } \dot{x}=0 \text{ 时的 } \cot\alpha = \frac{\dot{x}}{\dot{y}}\right)$ 有确定的值. 在构成 $\frac{d\alpha}{ds}$ 时，我们总可以假定，属于 P 点的一个邻域的参数值全都在这样一个区间上，在整个这个区间上 $\dot{x}$，$\dot{y}$ 之一恒不为零. 如果，譬如说，$\dot{x} \neq 0$，那么我们可以在整个区间上，给 α 指定一个对 t 连续变化的值

$$\alpha = \alpha(t) = \operatorname{arc}\tan\frac{\dot{y}}{\dot{x}} + n\pi,$$

其中 n 是固定的（可能是负的）整数，"arc tan"表示函数的主值，它在 $-\frac{\pi}{2}$ 和 $\frac{\pi}{2}$ 之间. 类似地，如果在这个区间上 $\dot{y} \neq 0$，我们可以对 α 取表示式[1]

$$\alpha(t) = \operatorname{arc}\cot\frac{\dot{x}}{\dot{y}} + n\pi = \frac{\pi}{2} - \operatorname{arc}\tan\frac{\dot{x}}{\dot{y}} + n\pi.$$

在任一情形，对任何一种参数表示，我们都通过直接微分得到

$$\dot{\alpha} = \frac{d\alpha}{dt} = \frac{\dot{x}\ddot{y} - \ddot{x}\dot{y}}{\dot{x}^2 + \dot{y}^2}.$$

由于还有（见第 377 页（12a））

1）我们可以定义 $\alpha(t)$ 为对**所有**的参数值连续的函数，办法是把整个的参数区间分成一些子区间，在每个子区间上，或者 $\dot{x} \neq 0$，或者 $\dot{y} \neq 0$. 在每个子区间上，我们可以用上面的表示式之一定义 $\alpha(t)$. 对于每个子区间，我们这样选择整常数 n，使得在两个相邻的区间的公共端点上，用表示式确定的 α 值相等.

$$\dot{s}=\frac{ds}{dt}=\sqrt{\dot{x}^2+\dot{y}^2},$$

我们得到曲线的曲率 $\frac{d\alpha}{ds}=\frac{\dot{\alpha}}{\dot{s}}$ 的表达式

$$\kappa=\frac{d\alpha}{ds}=\frac{\dot{\alpha}}{\dot{s}}=\frac{\dot{x}\ddot{y}-\dot{y}\ddot{x}}{(\dot{x}^2+\dot{y}^2)^{3/2}}. \tag{15}$$

特别地，选择弧长 s 作为参数 t，我们有

$$\dot{x}^2+\dot{y}^2=1$$

(见第 378 页，(13))．因此，我们得到简化了的结果

$$\kappa=\dot{x}\ddot{y}-\dot{y}\ddot{x}.$$

曲率的正负号和绝对值

引进新参数 $\tau=\tau(t)$ 代替 t 并不影响切线的方向，因之也不影响 α 的变化．类似地，两个点的 s 值之差的绝对值具有和参数的选择无关的几何意义，也就是沿曲线度量的距离．但是，差的正负号必须与相应的参数值之差的正负号一致，因为我们定义 s 为 t 的增函数．因此，曲率 $|\kappa|=\left|\frac{d\alpha}{ds}\right|$ 的绝对值不依赖于参数的选择，而 κ 的正负号依赖于曲线相应于 t 增加的指向．显然，$\kappa>0$

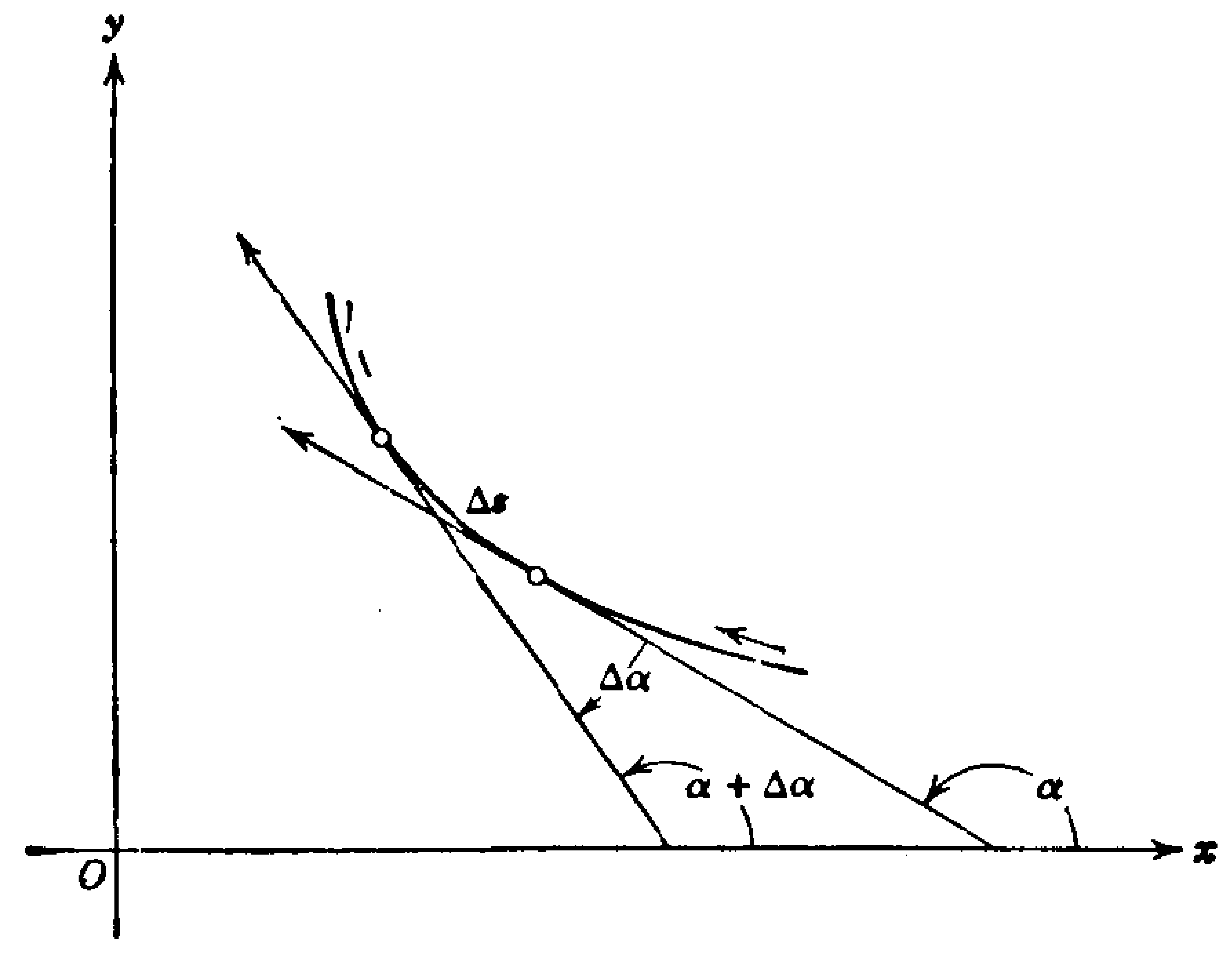


图 4.21a　曲线的曲率 $\kappa=\lim\frac{\Delta\alpha}{\Delta s}$ (图示的情况为 $\kappa<0$)

是指 α 随 s 增加而增加，即当我们沿着曲线的 s 或 t 增加的方向前进时，切线沿逆时针转动(见图 4.21a)．在这种情况下，曲线 C 的定向使得 C 的正侧也是 C 的内侧，即向着 C 弯曲的一侧．

如果曲线的方程为 $y=f(x)$，那么使用 x 作为参数，我们有

$$\kappa=\frac{y''}{(1+y'^2)^{3/2}}, \tag{16}$$

其中 y' 和 y'' 为 y 对变量 x 的导数．这里，曲率的符号是相应于 x 增加的符号．显然，对 $y''>0$，κ 是正的；在这种情况下，当 x 增加时，切线逆时针转动；我们称函数 $f(x)$ 是**下凸的**．连接任意两个点的曲线部分在连结这两点直线的下方．对 $y''<0$，当 x 增加时，切线顺时针转动，称函数是**上凸的**(图 4.22)．这里，曲线是在连结它的两个点的弦的上方．曲率的值为零的中间的情况(一般地说)相应于**拐点**，在拐点上，$y''=0$ (见第 251 页)．

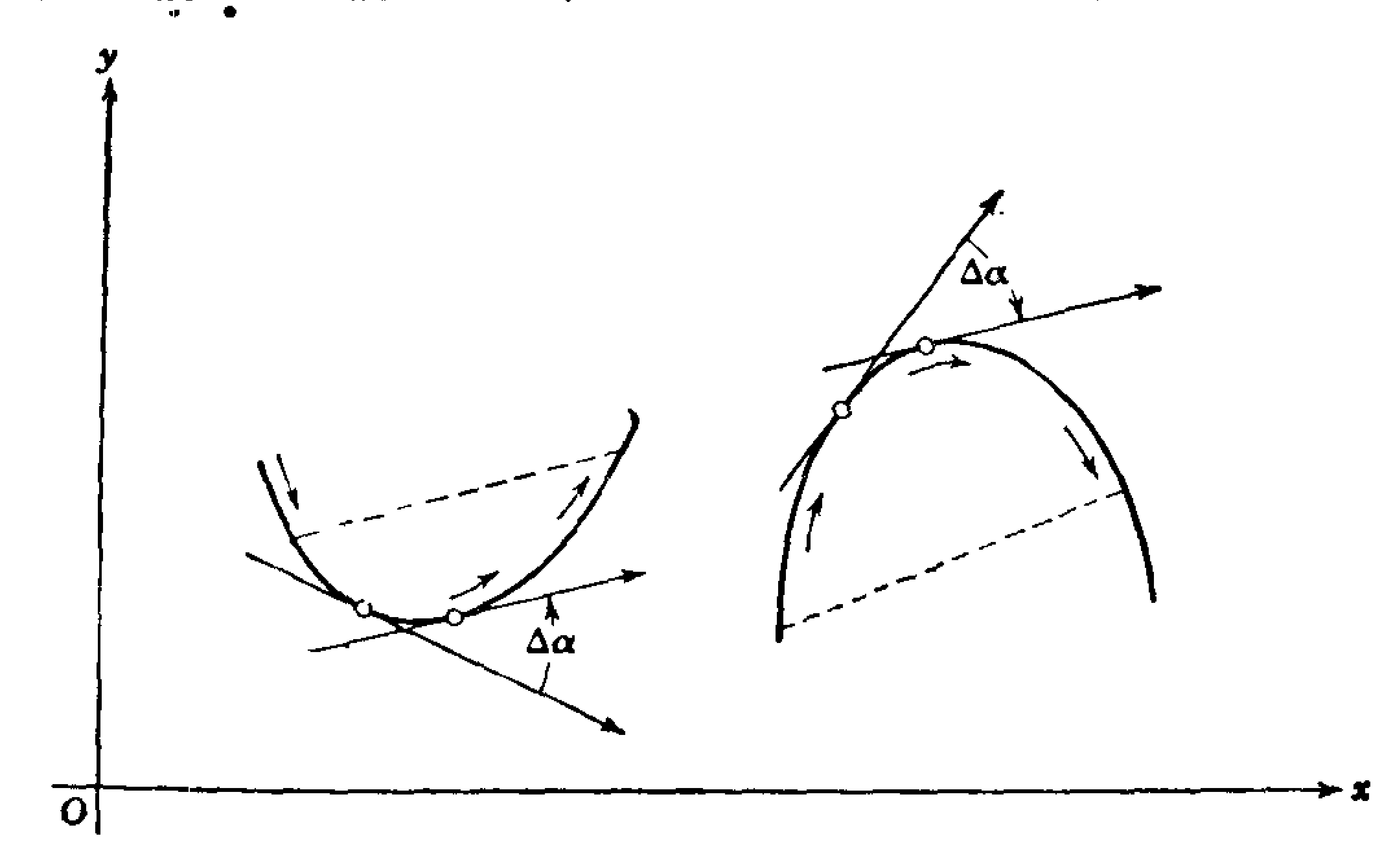


图 4.22 下凸函数的图形(左)和上凸函数的图形(右)

例

对于由方程 $x=a\cos t$，$y=a\sin t$ 给定的，半径为 a 的圆的曲率，从一般公式(15)，我们得到常数值 $\frac{1}{a}$．因此，以反时针的指向描绘的圆的曲率是半径的倒数．这个结果使我们确信，我们曲率的定义确实是恰当的定义；因为，在圆的情况下，我们很自然

地把半径的倒数看成它弯曲程度的一种测量.

第二个例子是函数 $y=x^3$ 定义的曲线. 曲率为

$$\kappa=\frac{6x}{(1+9x^4)^{3/2}}.$$

在 $x<0$ 时,由于 $\kappa<0$,所以函数 $y=x^3$ 是上凸的,并且切线顺时针转动;而在 $x=0$ 时,我们有一个拐点;在 $x>0$ 时,函数变为下凸的.

由定义容易看出,曲率恒等于零的函数是一条直线,而且只有直线,曲率才恒等于零.

曲率圆和曲率中心

我们引进 $\rho=\frac{1}{\kappa}$. 量 $|\rho|=\frac{1}{|\kappa|}$ 叫做在所讨论的点的**曲率半径**(在拐点上,$\kappa=0$,曲率半径为无穷大). 圆上任意一点的曲率半径就是圆的半径.

对于曲线 C 上任意一点 $P=(x,y)$,我们作一个在 P 与 C 相切的圆,使得我们在点 P 处以同样的指向通过曲线和圆的时候,该圆与 C 有相同的曲率.这个圆叫做曲线 C 在 P 点的**曲率圆**.它的中心

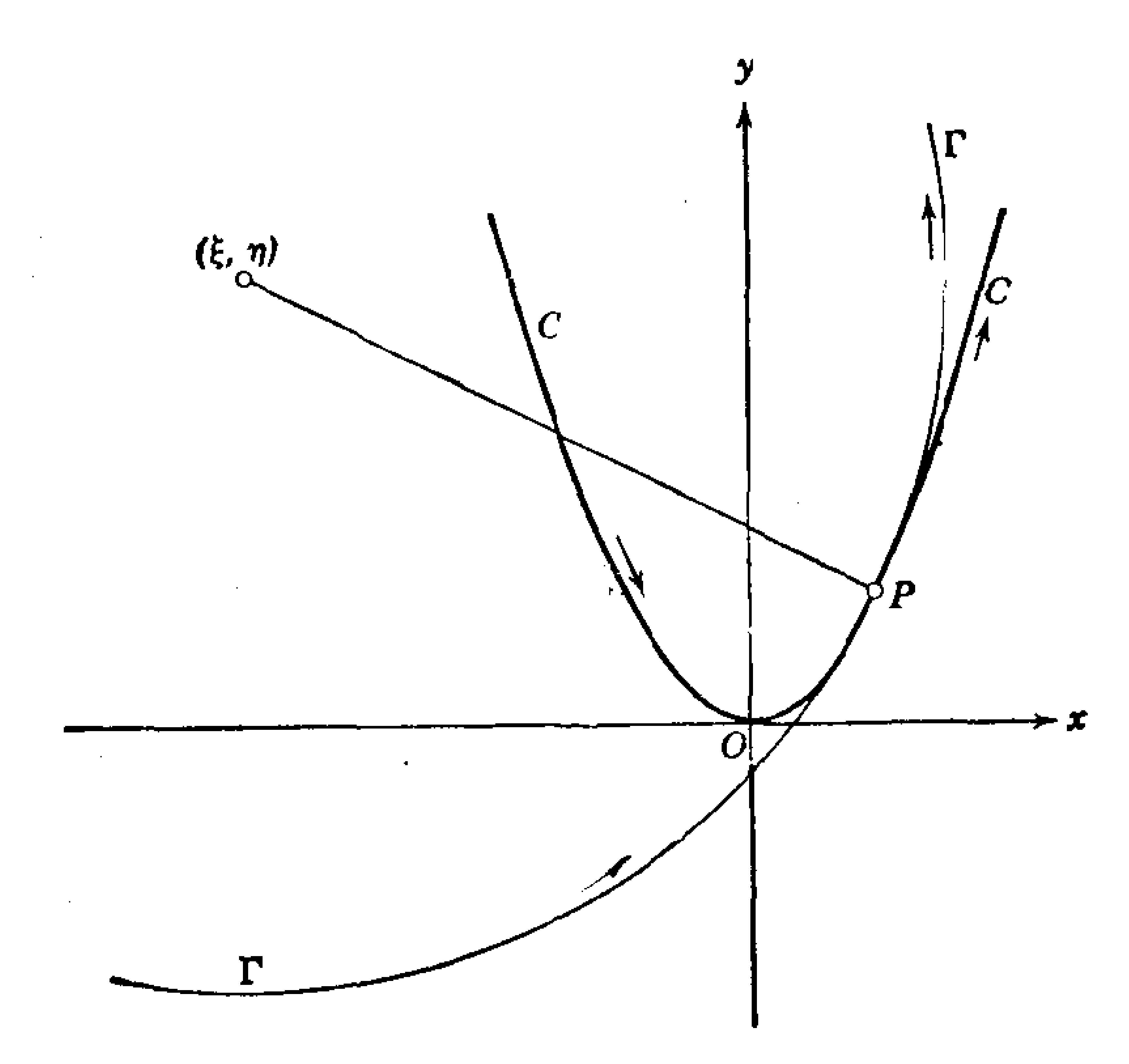


图 4.23 对应于曲线 C 的点 P 的曲率圆 Γ 和曲率中心 (ξ,η)

是曲线C相应于点P的曲率中心(图 4.23). 因为C和圆有相同的曲率半径,所以圆的半径必定是C的曲率半径$|\rho|$. 圆的中心(ξ, η)必定在C过P点的法线上,到P的距离是$|\rho|$. 因为C和圆向同一侧弯曲,所以中心位于曲线在点P的法向上,在正侧或在负侧根据曲率κ是正或负而定.

如果 $\kappa>0$, 那么由P到曲率中心的方向与正x轴夹角为 $\alpha+\frac{\pi}{2}$. 因此,如果ξ, η是曲率中心的坐标,x, y是点P的坐标,我们有(见第 369 页方程 (7))

$$\frac{\xi-x}{\rho}=\cos\left(\alpha+\frac{\pi}{2}\right)=-\sin\alpha=\frac{-\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}},$$

$$\frac{\eta-y}{\rho}=\sin\left(\alpha+\frac{\pi}{2}\right)=\cos\alpha=\frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}}.$$

因此,对 $\kappa>0$,

$$\xi=x-\frac{\rho\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}}, \quad \eta=y+\frac{\rho\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}}. \tag{17}$$

如果弧长s作为参数t,我们得到简单的表示式

$$\xi=x-\rho\dot{y}, \quad \eta=y+\rho\dot{x}. \tag{17a}$$

对 $\kappa<0$,得到ξ, η的同样的公式,这种情况下,曲率半径是$-\rho$,而且从P到中心的方向与正x轴的夹角为 $\alpha-\frac{\pi}{2}$.

曲率圆作为密切圆

公式(17)用曲线上点P的参数t给出了曲率中心的表示式. 当t取遍参数区间的所有值,曲线中心描出一条曲线,即所谓给定曲线的法包线 (evolute). 因为我们必须把$\dot{x}$, $\dot{y}$和ρ与x, y一起看成t的已知函数,所以前面的公式给出了这个法包线的参数方程. 法包线的例和几何性质的讨论将在附录 1 第 447 页中给出.

如果通过点P的任意两条曲线有相同的切线,并且当它们有相同的定向时还有相同的曲率,那么这两条曲线称为在点P密切或二阶相切. 显然,两条密切曲线在点P有相同的曲率圆和

曲率中心。如果两条曲线以非参数形式的方程 $y = f(x)$，$g = g(x)$ 给出，那么很容易表达它们有切点P，并且在点 P 有相同的切线和曲率的条件。如果 x 为切点 P 的横坐标，那么我们有 $f(x) = g(x)$，$f'(x) = g'(x)$；曲率相等表示为

$$\frac{f''(x)}{[1 + f'^2(x)]^{3/2}} = \frac{g''(x)}{[1 + g'^2(x)]^{3/2}},$$

因此 $f''(x) = g''(x)$。这样，具有相等曲率的切点的条件是在这点上 f 和 g 的值及其一阶、二阶导数的值都相等。

考虑曲线 C：$y = f(x)$ 和它在点 P 的曲率圆 Γ，圆 Γ 在 P 点的一个邻域内表示为 $y = g(x)$。因为圆 Γ 与它自己的曲率圆重合，所以我们看到 C 和 Γ 有相同的曲率圆，因之，在 P 点 C 和 Γ 密切。所以在切点上有 $f(x) = g(x)$，$f'(x) = g'(x)$，$f''(x) = g''(x)$。我们说这个圆在切点 P 对于曲线是最佳拟合的圆，因为任何其他在切点与曲线相交的圆都不和 C 在该点"二阶相切"。曲率圆就是密切圆(参看第五章第 485 页)。

附带说一说，犹如曲线 C 的切线是过 C 上两相邻点 P 和 P_1 的直线当 $P_1 \to P$ 时的极限一样，我们能够证明，在 P 点的曲率圆是通过三点 P, P_1, P_2 的圆当 $P_1 \to P$ 和 $P_2 \to P$ 时的极限。证明留给读者(见第 461 页问题 4)

i. 坐标轴变换，不变量

几何或物理状况所固有的一些性质不依赖用以表述它们的特殊的坐标系或"参考标架"；像距离、长度或角度这些性质是固有的，这一点必须反映在一些关系式中，这些关系式表明，当我们由一个坐标系过渡到另一个坐标系时，有关的各个公式保持不变或者说是**不变量**(invariance)。关于这个问题的几个简单的说明在这一节叙述是合适的。

我们使用普遍方程把一个坐标系中点 P 的坐标 x, y 和任何另一坐标系中同一点 P 的坐标 ξ, η 联系起来。第二组坐标对于第一组坐标的相对位置可以用第二坐标系的原点在第一坐标系的坐标

a, b 和正 ξ 轴与正 x 轴的夹角 γ 来表示.[1]同一个点在两个坐标系中的坐标 (x, y) 和 (ξ, η) 是通过变换

$$\begin{aligned} x &= \xi\cos\gamma - \eta\sin\gamma + a, \\ y &= \xi\sin\gamma + \eta\cos\gamma + b \end{aligned} \tag{18}$$

联系着的（参见图 4.24）．对于 $\gamma=0$，只包含平行位移，即平移(translation)，而没有轴的旋转．公式简化为 $x=\xi+a$，$y=\eta+b$.

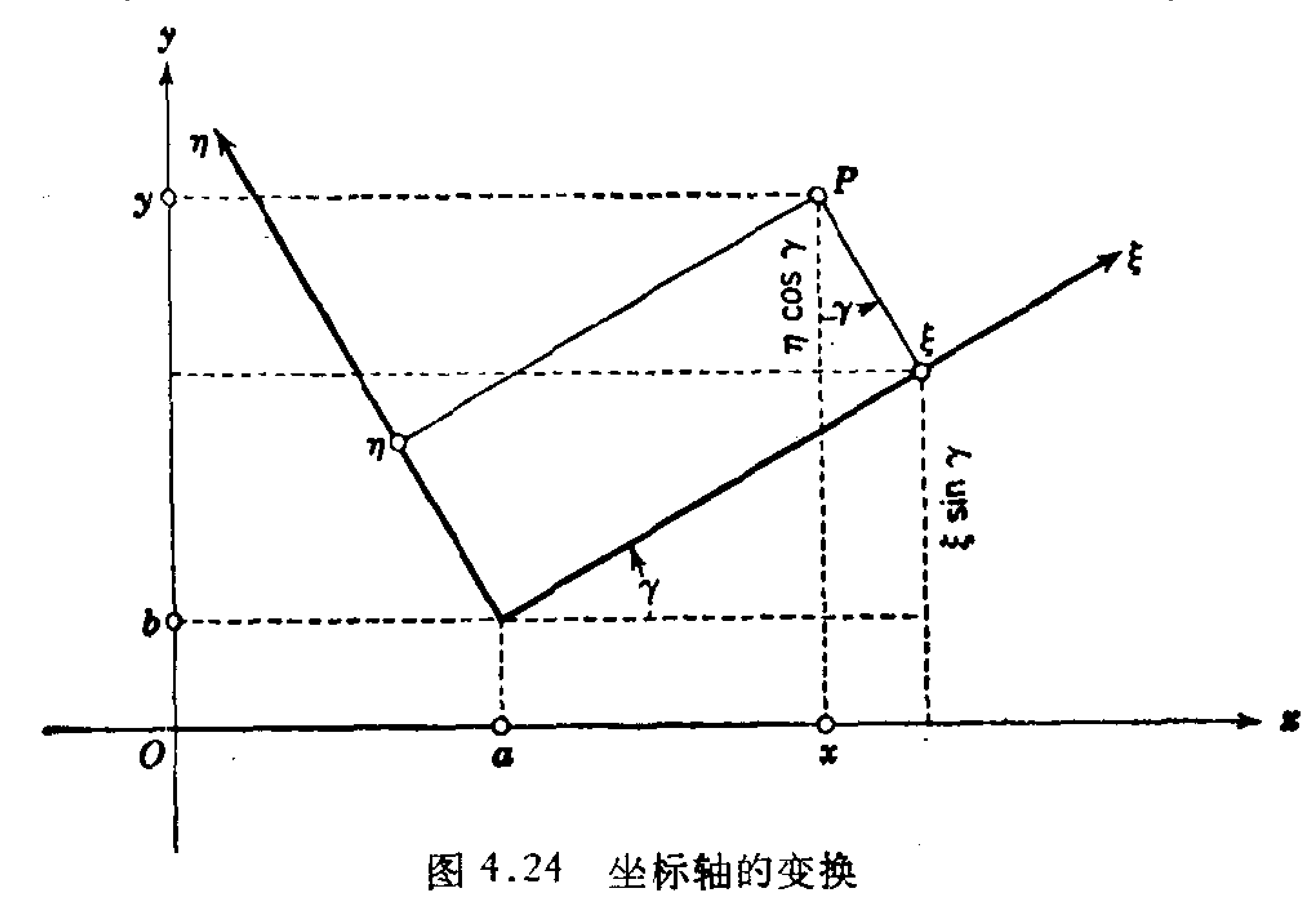


图 4.24　坐标轴的变换

由 x, y 解出 ξ, η，我们有

$$\begin{aligned} \xi &= (x-a)\cos\gamma + (y-b)\sin\gamma, \\ \eta &= -(x-a)\sin\gamma + (y-b)\cos\gamma. \end{aligned} \tag{18a}$$

如果 x 和 y 是确定曲线的参数 t 的函数，那么根据这些公式我们立即得到 ξ 和 η 作为 t 的函数的表示式，它们给出了同一曲线在 ξ, η 坐标系中的参数表示．对 t 微分（确定两个坐标系的相对位置的量 a，b，γ 不依赖于 t），就得到"速度分量"（即坐标对 t 的导数）的变换[2]，

1) 我们只限于右手坐标系，即坐标系第二个轴的正方向是由第一个轴的正方向经反时针旋转 90° 而得到的．

2) 在某些物理应用中，t 代表着时间，两个坐标系的相对位置也依赖于时间．设量 x，y 表示在静止的坐标系中一个质点的坐标，而 ξ, η 是同一个质点相对于运动着的坐标系的坐标，例如坐标轴放在运动的地球上．函数 $x(t)$，$y(t)$ 描绘了静止的观察者看到的质点的路径，而 $\xi(t)$，$\eta(t)$ 描绘了运动着的观察者看到的路径．那么，$\dot{x}$，$\dot{y}$ 和 $\dot{\xi}$，$\dot{\eta}$ 相联系的公式必须也包含从微分 a，b，γ 而得出的各项．

$$\dot{x} = \dot{\xi}\cos\gamma - \dot{\eta}\sin\gamma,$$
$$\dot{y} = \dot{\xi}\sin\gamma + \dot{\eta}\cos\gamma.$$

我们证实了

$$\dot{x}^2 + \dot{y}^2 = \dot{\xi}^2 + \dot{\eta}^2.$$

因此，表达式 $\sqrt{\dot{x}^2 + \dot{y}^2}$ 在所有的坐标系中有相同的值. 当然，把这个量理解为沿曲线的长度对时间 t 的变化率 $\frac{ds}{dt}$，它的不变性就显而易见了. 读者可以通过简单的计算证明曲率公式 $\kappa = (\dot{x}\ddot{y} - \ddot{x}\dot{y})(\dot{x}^2 + \dot{y}^2)^{-\frac{3}{2}}$ 也是不变的. (当然这一点也可由下述事实直接推出，即切线与 ξ 轴和 x 轴的两夹角仅相差一个常数值 γ，因之 $\kappa = \frac{d\alpha}{ds}$ 不能改变.)

联系坐标 x，y 与坐标 ξ，η 的方程(18)常常另外理解为描写位移. 在这种理解中，改变了的是点 P，而不是坐标轴(图 4.25). 只用一个坐标系，在那个坐标系中坐标为 (x, y) 的点映射到同一坐标系中坐标为 (ξ, η) 的点. 曲线的长度和曲率的不变性现在的含义是，当整个曲线作刚体运动时这些量不改变.

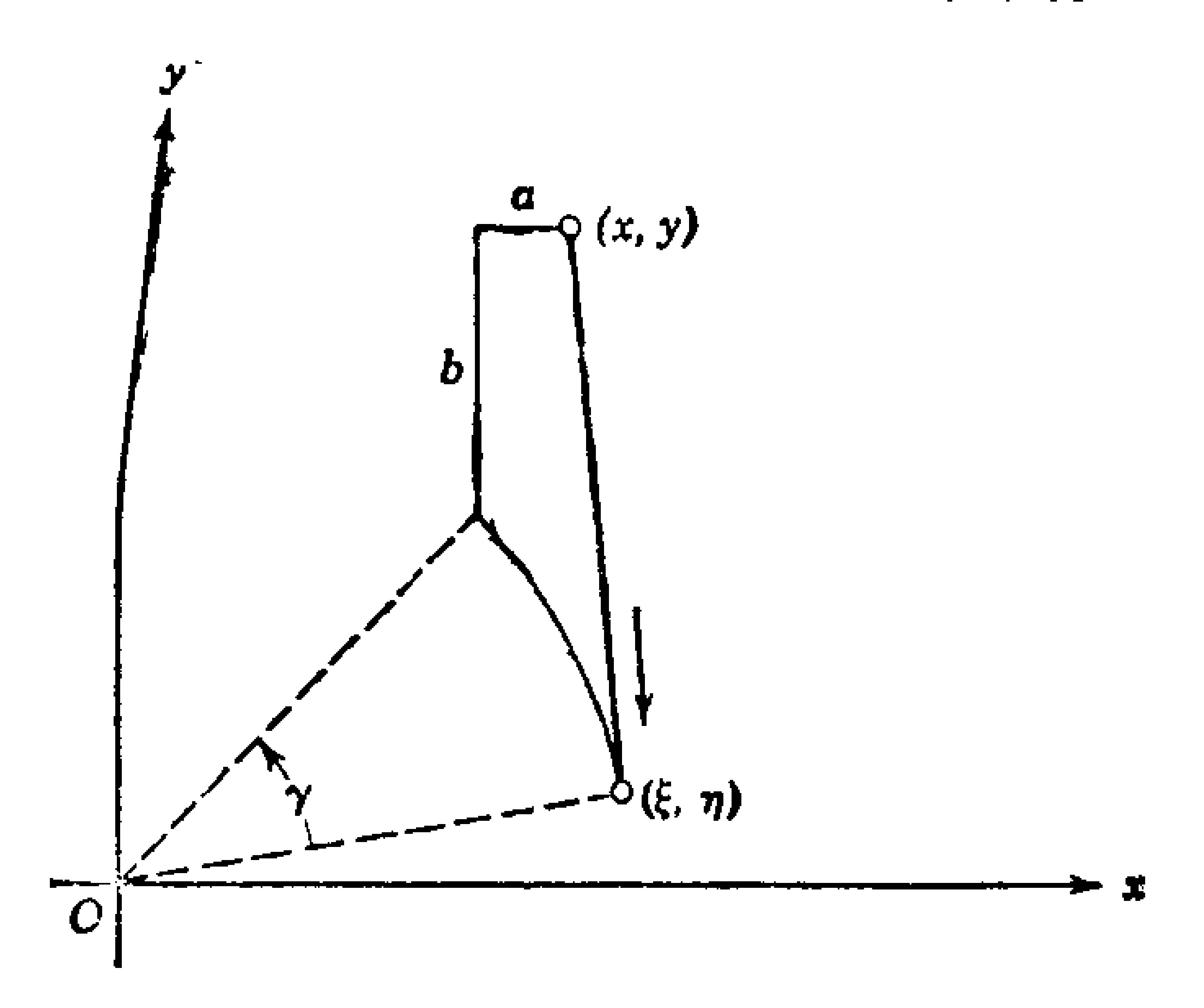


图 4.25　点 P 从位置 (x, y) 到位置 (ξ, η) 的位移

*j. 狭义相对论中的匀速运动

在第 242 页已经指出，在三角函数和双曲函数之间有一些很

深刻的类似点，它们在几何方面的表现就是椭圆与双曲线的性质之间的对应关系．当我们能够定义虚变量的三角函数并且能够证明在第 7.7a 节中的 $\cos(it)=\cosh t$，$\sin(it)=i\sinh t$ 时，两者的关系就变得清楚了．作为这个相似点的一种应用，我们考虑一个平面的“双曲旋转”，它可以与爱因斯坦（Einstein）的狭义相对论中直线的洛伦兹（Lorentz）变换等同起来．

在第 385 页(18a)式中，我们看到原点保持固定，坐标轴旋转一个角度 γ 的旋转，可以用方程组

$$\begin{aligned}\xi &= x\cos\gamma + y\sin\gamma,\\ \eta &= -x\sin\gamma + y\cos\gamma\end{aligned}\tag{18b}$$

来描述，这把点 P 在第一个坐标系的坐标和它在第二个坐标系的坐标 ξ,η 联系起来．从原点到 P 的距离在两个坐标系中有同样的表示式：

$$OP=\sqrt{x^2+y^2}=\sqrt{\xi^2+\eta^2}.$$

如果我们利用恒等式 $\cos^2\gamma+\sin^2\gamma=1$，这也可由变换方程直接推出．

现在我们考虑系数为双曲函数，而不是三角函数的类似的变换：

$$\begin{aligned}\xi &= x\cosh\alpha - t\sinh\alpha,\\ \tau &= -x\sinh\alpha + t\cosh\alpha;\end{aligned}\tag{19}$$

在旋转公式(18b)中把旋转角 γ 和 y 与 η 的坐标取为纯虚量：$\gamma=i\alpha$，$y=it$，$\eta=i\tau$，就得到上式．

我们注意，若 α 为实值(其意义就是在原来的理解中旋转角 γ 是虚数)，式(19)定义 ξ 和 τ 为 x 和 t 的实线性函数．这些函数有下列特殊的性质

$$\begin{aligned}\xi^2-\eta^2 &= (x\cosh\alpha - t\sinh\alpha)^2-(-x\sinh\alpha+t\cosh\alpha)^2\\ &= x^2-t^2.\end{aligned}$$

这是恒等式 $\cosh^2\alpha-\sinh^2\alpha=1$ 的结果．(当然，这也可以从观察到 $x^2-t^2=x^2+y^2$ 是 x,y 平面上到原点距离的平方而得到．)现在，我们把 t 理解为时间，把 x 理解为描述一维空间，即一条直

线上点的位置的空间坐标. 任何一个事件都是发生在某一时刻的某一地点的. 这两个信息由 x, t 给出, x 为从原点 O 到该点的距离(带正负号), t 为由时刻 0 开始经过的时间. 在相对论中, 我们的观点是, 这个距离和经过的时间的测得值依赖于观察者使用的参考系,即时空连续统的特殊坐标系. 式(19)得到的量 ξ, τ 描述了在不同的参考系中的同一个事件, 在不同的参考系中距离和时间区间的长度可以取不同的值. 从一个参考系到另一个参考系的变换(即熟知的洛伦兹变换)中, 不变的量是原点到事件的"时空距离"

$$\sqrt{x^2 - t^2} = \sqrt{\xi^2 - \tau^2}.$$

对于使用第二个坐标系的观察者,量 ξ 是由原点 $\xi = 0$ 度量的空间距离. 原点是这样的点:

$$x \cosh \alpha - t \sinh \alpha = 0,$$

即 $\frac{x}{t} = \tanh \alpha$. 因此,第二个坐标系的原点在第一个坐标系中是一个相对于该坐标系的原点以匀速 $v = \frac{dx}{dt} = \tanh \alpha$ 运动的点. 因此, 洛伦兹变换把相互以恒速运动的两个系统上的观察者面前呈现的距离和时间的值联系了起来. 这里,

$$v = \frac{\sinh \alpha}{\cosh \alpha} = \frac{e^{\alpha} - e^{-\alpha}}{e^{\alpha} + e^{-\alpha}}$$

必须介于 -1 和 $+1$ 之间,所以我们所讨论的只限于两个系统的相对速度在数值上小于 1 的情况. 这里, 数值 1 表示不能被任何速度 v 超过的光速 c, 它的单位是适当选择的.

对于一个常数 u, 方程 $x = ut$ 相应于在第一个系统中从 $x = 0$ 和时刻 $t = 0$ 开始以速度 u 运动的点. 这同一个点在第二个系统中的速度为

$$\omega = \frac{d\xi}{d\tau} = \frac{d\xi}{dt} \Big/ \frac{d\tau}{dt} = \frac{u - \tanh \alpha}{1 - u \tanh \alpha} = \frac{u - v}{1 - uv}.$$

在爱因斯坦狭义相对论中成立的这个结果与我们在古典运动学中得到的结果不同, 在古典运动学中, 对于以速度 v 运动着的系统,

一个点的相对速度ω由$\omega = u - v$简单地给出．相对论的公式表明，当$u = +1$或-1时$\omega = u$．这相应于由著名的迈克尔逊-莫雷（Michelson-Morley）实验推断的事实，即对以不同的速度运动的观察者来说，光的速度是不变的．

k．表示闭曲线内部面积的积分

在第二章中，参照"曲线下的面积"，即特殊形状的长条的面积，提出了积分的概念．这样限定于曲线下的面积是不十分令人满意的，因为我们实际上遇到最多的是闭曲线C内部区域的面积，它们有比用积分$\int_a^b f(x)dx$表示面积的条形更一般的形状．

基本公式

现在，假定一条简单闭曲线C用参数表示给出，我们将推导C所围面积的一个漂亮的普遍积分表示式，办法是把该面积分成特殊的条形面积．这个表示式将不依赖于参数表示，也不依赖于坐标系．而且它还表示根据边界C的指向的曲线内部的有向面积；就是说，按边界曲线C的指向是顺时针还是逆时针来指定简单闭曲线中的面积是负号还是正号．

假定有向简单闭曲线C由$x = x(t)$，$y = y(t)$给出，其中t在区间$\alpha \leqslant t \leqslant \beta$上变化，$t$增加的指向确定为$C$的指向．我们假定$x$和$y$是$t$的连续函数（在$t = \alpha$和$t = \beta$时有同样的值）并且它们的一阶导数$\dot{x}$和$\dot{y}$是连续的，除了当$C$有角点时可能有有限个跳跃不连续点之外．在这些假定之下，我们将证明在C内的有向面积A的这一基本公式

$$A = -\int_\alpha^\beta y\dot{x}dt = \int_\alpha^\beta x\dot{y}\,dt = \frac{1}{2}\int_\alpha^\beta (x\dot{y} - y\dot{x})dt. \qquad (20)$$

对第一个积分使用分部积分法，利用周期条件$x(\alpha) = x(\beta)$，$y(\alpha) = y(\beta)$，我们可直接推出公式中三个积分表示式是等价的；第三个更对称的表示式刚好是前两个的算术平均值．

公式（20）不依赖平面上坐标系的位置．事实上，对称表示式

$$A=\frac{1}{2}\int_{\alpha}^{\beta}(x\dot{y}-y\dot{x})dt$$

清楚地表明A的值不依赖于坐标系的选择. 如我们在第 385 页看到的, x, y 坐标系到 ξ, η 坐标系的变换由替换

$$x=\xi\cos\gamma-\eta\sin\gamma+a,$$
$$y=\xi\sin\gamma+\eta\cos\gamma+b$$

实现,其中 a, b, γ 为常数. 把这些公式对 t 微分,就有

$$\dot{x}=\dot{\xi}\cos\gamma-\dot{\eta}\sin\gamma,$$
$$\dot{y}=\dot{\xi}\sin\gamma+\dot{\eta}\cos\gamma,$$

所以

$$x\dot{y}-y\dot{x}=\xi\dot{\eta}-\eta\dot{\xi}+a\dot{y}-b\dot{x}.$$

因此,在绕原点旋转(即当 $a=b=0$ 时)之下 $x\dot{y}-y\dot{x}$ 是不变量. 即使 a 或 b 不为零, A 的积分值也不受影响,因为对于闭曲线 C 有

$$\int_{\alpha}^{\beta}(a\dot{y}-b\dot{x})dt=(ay-bx)\Big|_{\alpha}^{\beta}=0.$$

基本公式(20)的证明. 在简单弧上的曲线积分

我们分几个平易的步骤来证明基本公式(20).

首先, 设C是一个有起点 P_0 和终点 P_1 的简单的有向弧. 设 $x=x(t)$, $y=y(t)$ 是 C 的任一参数表示, P_0, P_1 分别对应于 $t=t_0, t_1$ (这里 t_0 可能大于 t_1, 也可能小于 t_1). 那么积分

$$A=-\int_{t_0}^{t_1}y\frac{dx}{dt}dt$$

仅依赖于C,而不依赖特殊的参数表示. 这是换元法的显然推论; 如果我们用单调函数 $\tau=\chi(t)$ 引人新的参数 τ, 其中 $\tau_0=\chi(t_0)$, $\tau_1=\chi(t_1)$, 那么相应的积分是[1]

$$-\int_{\tau_0}^{\tau_1}y\frac{dx}{d\tau}d\tau=-\int_{t_0}^{t_1}y\frac{dx}{d\tau}\frac{d\tau}{dt}dt=-\int_{t_0}^{t_1}y\frac{dx}{dt}dt=A.$$

---

1) 我们假定不仅 $x(t)$, $y(t)$ 是连续函数,而且 $\tau(t)$ 也是连续函数, 并且它们的导数是连续的,除了可能有的有限个跳跃不连续点之外.

所以有理由在积分 A 的表达式中去掉任何特殊的参数 t 的关系而简记为

$$A = A_C = -\int_C y dx.$$

这里，对于有向简单弧 C 的 A_C 按如下方法计算，把弧联系于参数 t，利用 $dx = \left(\frac{dx}{dt}\right)dt$，把 C 的端点的参数值按 C 的定向所决定的次序取为积分限.[1]

如果 C' 是把 C 改变定向而得到的弧，即起点为 P_1，终点为 P_0 的弧，那么对 C' 用同一个参数表示，我们有

$$A_{C'} = -\int_{t_1}^{t_0} y \frac{dx}{dt} dt = +\int_{t_0}^{t_1} y \frac{dx}{dt} dt = -A_C.$$

因此，如改变弧 C 的定向，那么积分 A_C 的正负号也要改变.

如果有向简单弧 C 分成每一个都与 C 有相同定向的有向子弧 $C_1, C_2, \cdots, C_n$，那么我们显然有

$$A_C = A_{C_1} + A_{C_2} + \cdots + A_{C_n}.$$

因为在 C 的一个参数表示中，譬如说 C 的指向是 t 增加的指向，那么这个分解相应于把 C 的参数区间 $t_0 \leqslant t \leqslant t_n$ 剖分成对应于 $C_1, \cdots, C_n$ 的子区间 $t_0 \leqslant t \leqslant t_1, t_1 \leqslant t \leqslant t_2, \cdots, t_{n-1} \leqslant t_n$. 由积分的可加性就推出上述结果.

当 C 由几段光滑的弧 $C_1, C_2, \cdots$ 组成，它们中每一个都有自己的参数表示，积分 A_C 的可加性使计算 A_C 的值更容易. 我们不必要人为地构造整个曲线 C 的一般的参数表示，而只要从每一个参数表示计算 A_{C_i}，然后取和. 而且 A_{C_i} 可以以任意次序相加；我们只须保证所有的 C_i 与 C 有相同的定向.

闭曲线的基本曲线积分

现在，我们能够对任意有向简单闭曲线 C 定义 A_C，办法是把

---

1) 积分 $\int_C y dx$ 是一般**曲线积分** $\int_C P dx + Q dy$ 的一个例子，我们将在第二卷中讨论它.

C 分成与 C 有一致的定向的简单弧 C_1, $\cdots$, C_n，然后取 A_{C_i} 的和.[1]如果整个闭曲线 C 有参数表示 $x=x(t)$, $y=y(t)$, $\alpha\leqslant t\leqslant\beta$, 其中 C 的定向为 t 增加的指向，并且 $t=\alpha$ 和 $t=\beta$ 对应于同一个点，那么仍然有

$$A_C=-\int_\alpha^\beta y\frac{dx}{dt}dt.$$

我们可用同样方法对分解为有向简单弧的非简单的有向曲线 C 定义 A_C, 即使 C 是由几个不相交的部分组成的也可以这样作，只要 C 的每一个部分有确定的指向.

基本积分作为面积

现在我们转向主要之点，即闭曲线的表达式 A_C, 就是在 C 内的有向面积这样一个直观的几何量.

我们首先考虑以弧 C_1: $y=g(x)$, $a\leqslant x\leqslant b$ 为上界；以弧 C_3: $y=f(x)$, $a\leqslant x\leqslant b$ 为下界；侧边为由 $x=a$ 和 $x=b$ 给出的线段 C_2, C_4 的区域 G(图 4.25(a)). 这里，C_2 和 C_4 允许缩成为一点. 如果我们给 C 以逆时针定向，弧 C_1 将以 x 减小的指向描

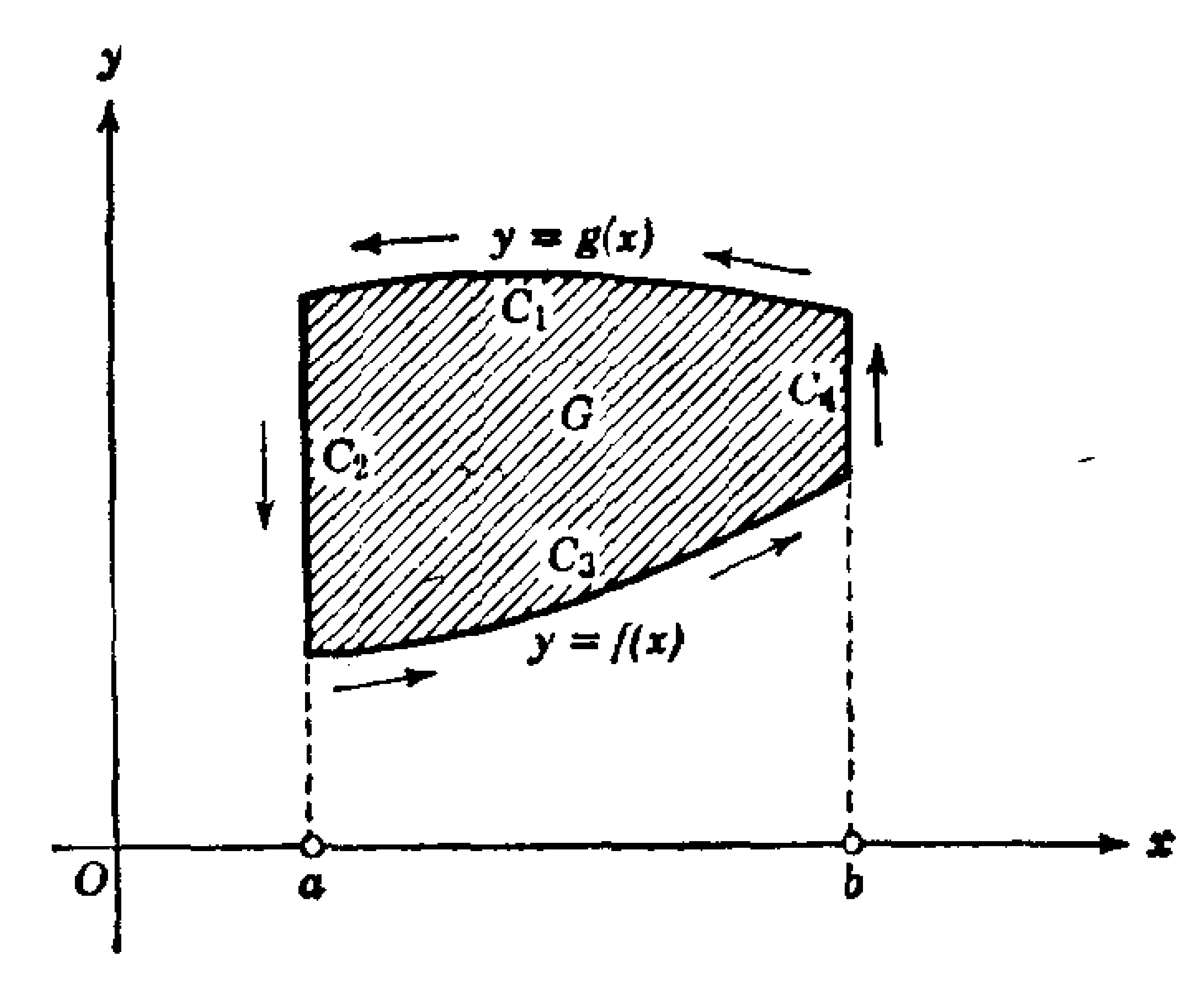


图 4.25(a) 一个单元的面积

1) 容易推论，以这种方式得到的 A_C 的值不依赖于我们把 C 分成简单弧的特殊分法：首先，关于简单弧的 A 的可加性表明，把一个给定的剖分用引入另外的剖分点而加密并不改变 A_C 的结果值；进一步，任意两个剖分可以用把两个都加密了的一个剖分来代替，而不改变 A_C 的结果值.

绘出来，而弧 C_3 则以 x 增加的指向描绘出来．把 A_C 取为四个 A_{C_i} 的和，x 为常数的 C_2 和 C_4 没有任何贡献，因为 $\frac{dx}{dt}=0$．在弧 C_1 和 C_3 上，用 x 作为参数，我们有

$$A_C = A_{C_1} + A_{C_3} = -\int_b^a g(x)dx - \int_a^b f(x)dx$$
$$= \int_a^b g(x)dx - \int_a^b f(x)dx.$$

当 G 完全在 x 轴的上方时，A_C 显然是区域 G 的正面积，它等于曲线 C_1 和 C_3 下面的面积之差．我们总可以保证 G 位于 x 轴上方，办法是用 $y+C$ 代替 y，C 为一适当的常数，也就是用 y 方向的一个变换．象我们前面看到的，这样作并不改变面积，也不影响闭曲线的 $A_C=-\int_C ydx$ 的值．因此，如果曲域 G 为上述类型，其边界 C 与平行于 y 轴的直线最多有两个交点，那么积分 A_C 表示面积．取正号或负号是根据 C 有逆时针定向或顺时针定向而定．如果曲线 C 与平行于 x 轴的直线最多有两个交点，那么，我们对 C 内的面积就得到同样的结果，只要把 A_C 记为 $\int_C xdy$，并在上述讨论中交换 x 和 y 的位置．我们把这样的两种类型之一的区域 G 称为“单元”．当它们的边界曲线给定这个或那个定向时，我们就说它们是“有向单元”．

我们现在考虑有向边界为 C 的区域 G，它由边界分别为 C_1，$C_2,\cdots,C_n$ 的若干个简单单元 G_1，$G_2,\cdots$，G_n 所组成；我们假定所有这些单元有同样的定向，譬如是逆时针的．两个相邻单元有部分公共边界，把这个公共边界分别考虑为两个单元的边界弧，则它们描绘了不同的指向，如图4.26所示．所以，如果我们把不同单元的积分 $A_{C_i}=-\int_{C_i} ydx$ 相加，那么所有内部单元边界的贡献都互相抵消，我们得到

$$A=\sum_{i=1}^{n} A_{C_i}=\sum_{i=1}^{n}\left(-\int_{C_i} ydx\right)=-\int_C ydx=A_C,$$

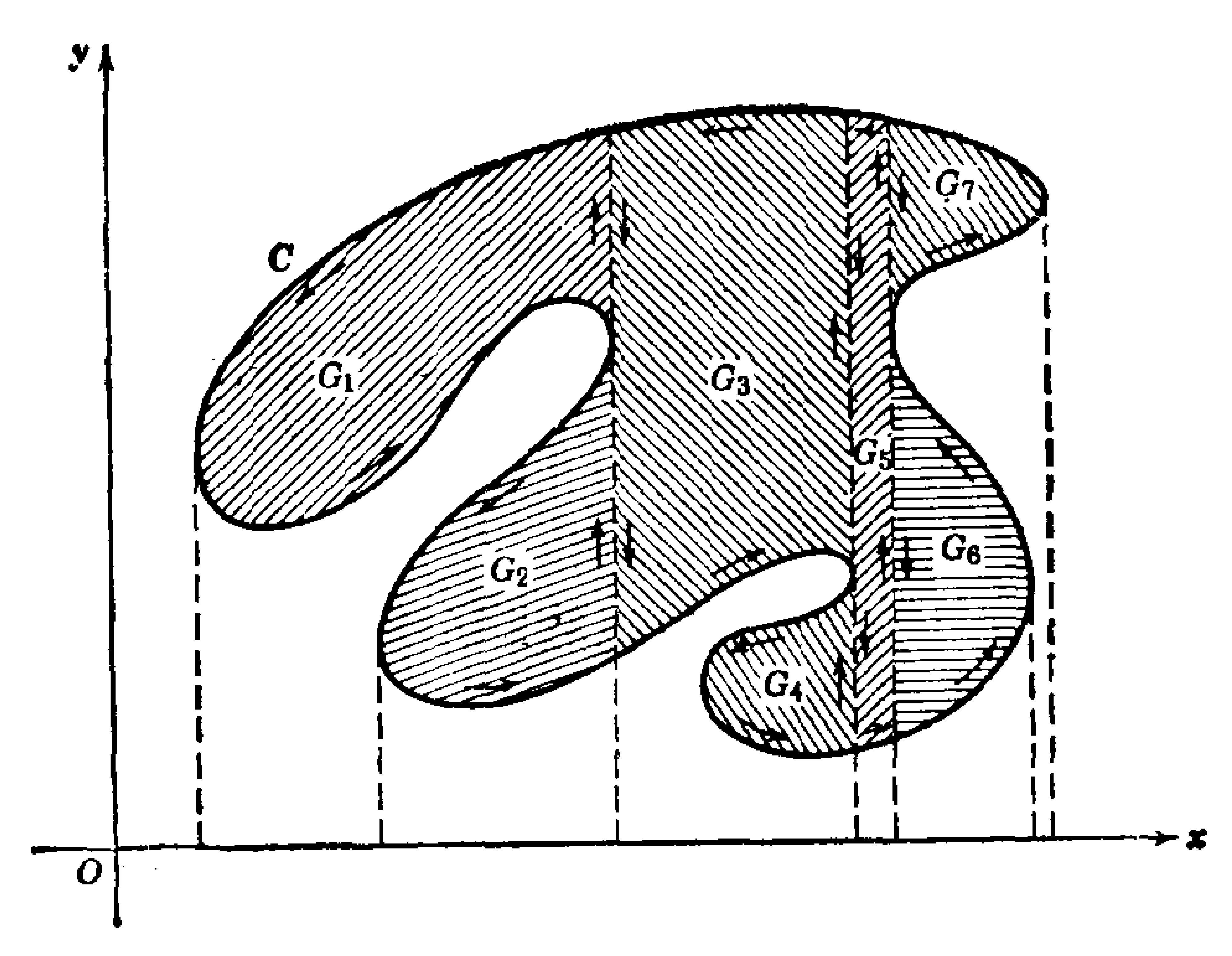


图 4.26 有向区域分解为有向单元

其中A是整个区域G的有向面积.

因此，我们证明了闭曲线内部的有向区域G的面积A的公式(20)，只要这个区域能分解成有限个简单单元，例如，可以用坐标轴的平行线进行分解.

对于我们将遇到的所有的区域，这个假定都是显然满足的，例如多边形区域.

补充说明

最后，附带说明，即使是多连通区域，如像环形域，只要它能分解成有限个简单单元，用同样方法可以证明面积公式的正确性. 所有的边界曲线的指向必须一致，即区域G的内部或者总是在左边或者总是在右边.

即使C不是简单曲线，而允许和自己相交，它把平面分成两个以上区域，A的公式仍然是有意义的. 在这种情况下，我们可以把公式看成为出发点，而把面积适当地理解为以C为界的平面上各连通小块的有向面积的加性组合. 我们将在本章附录II中讨论这个问题.

例. 作为例子，我们求椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 所包围的面积. 我

们取椭圆逆时针定向，根据参数表示 $x = a\cos t$，$y = b\sin t$，$0 \leqslant t \leqslant 2\pi$，我们有

$$A = \frac{1}{2}\int_0^{2\pi} (x\dot{y} - y\dot{x})dt = \frac{1}{2}\int_0^{2\pi} ab\,dt = \pi ab.$$

极坐标中的面积. 为使用极坐标 r，θ 表示面积，我们首先考虑以曲线段 $r = f(\theta)$ 和射线 $\theta = \alpha$，$\theta = \beta$ 为界的区域的面积 A. 假定 $\alpha < \beta$，并且 θ 能够作为沿曲线的参数（即不同的点有不同的极角）. 我们使用 A 的公式

$$A = \frac{1}{2}\int (x dy - y dx) = \frac{1}{2}\int (x\dot{y} - y\dot{x})dt,$$

这个积分必须分布在边界的曲线部分和两条射线上. 在射线 $\theta = \alpha$ 和 $\theta = \beta$ 上，我们可以用 r 作为参数. 由 $x = r\cos\theta$，$y = r\sin\theta$ 和 $\theta =$ 常数，我们有 $\dot{x} = \cos\theta$，$\dot{y} = \sin\theta$，因此 $x\dot{y} - y\dot{x} = 0$. 在曲线部分，我们用 θ 作为参数，则

$$\dot{x} = \frac{dr}{d\theta}\cos\theta - r\sin\theta,$$

$$\dot{y} = \frac{dr}{d\theta}\sin\theta + r\cos\theta.$$

因此 $x\dot{y} - y\dot{x} = r^2$. 所以

$$A = \frac{1}{2}\int_\alpha^\beta r^2 d\theta = \frac{1}{2}\int_\alpha^\beta f^2(\theta)d\theta. \tag{21}$$

对于简单闭曲线 C，假定 C 包含原点在它的内部并且 C 与原点发出的每一条射线恰交于一点，我们可以把 θ 作为参数，其中 $0 \leqslant \theta \leqslant 2\pi$. 那么被包围的面积是

$$A = \frac{1}{2}\int_0^{2\pi} r^2 d\theta. \tag{22}$$

极坐标面积公式 (21) 也可直接由积分的定义推导. 为此目的，我们把区域用从原点引出的射线分成扇形（图 4.27）. 每一个扇形由不等式

$$\theta_{i-1} < \theta < \theta_i, \quad 0 < r < f(\theta)$$

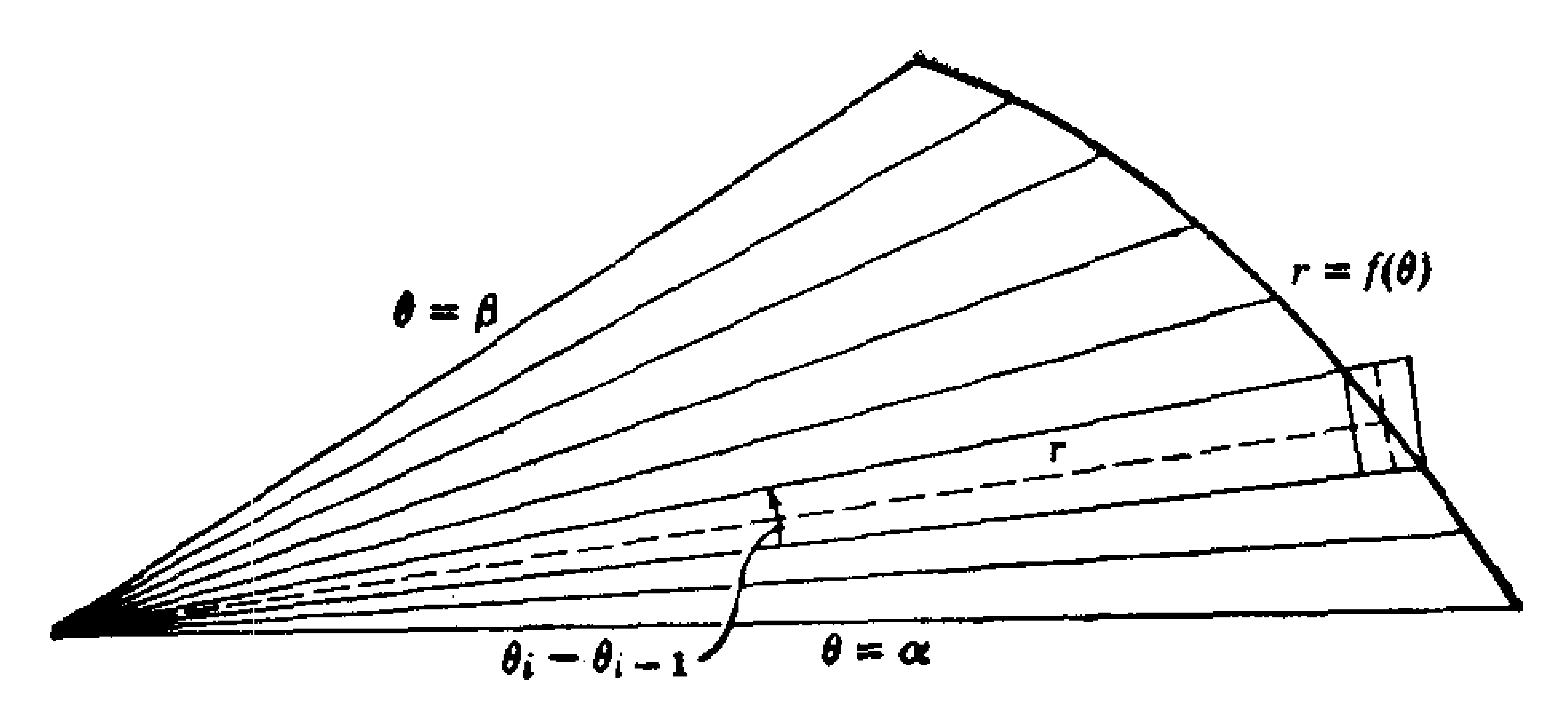


图 4.27 极坐标中的面积

描述，显然，扇形面积介于内接圆扇形和外接圆扇形面积之间；因此区域的一个扇形的面积等于 $\frac{1}{2}r^2(\theta_i-\theta_{i-1})$，其中 r 介于 $f(\theta)$ 在区间 $\theta_{i-1}<\theta<\theta_i$ 上的最大值和最小值之间. 当我们把剖分加密时，我们的区域的扇形面积之和显然收敛于积分 $\frac{1}{2}\int_\alpha^\beta r^2d\theta$.

双纽线的面积

作为方程(21)的例子，我们考虑由双纽线的一个回路包围的面积. 双纽线的方程(参见第 109 页)是 $r^2=2a^2\cos 2\theta$；θ 由 $-\frac{\pi}{4}$ 变到 $\frac{\pi}{4}$，就得到一个回路. 因此，对于该面积，我们有

$$a^2\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}}\cos 2\theta d\theta=a^2.$$

当然，另一个回路面积有相同的绝对值，但要取负值.

双曲线包围的面积

我们现在考虑以双曲线 $x^2-y^2=1$ 为界的扇形面积，在第 249 页中我们用了相当麻烦的方法计算过它(见图 3.12). 对于双曲线(更确切地说是对于它的右边的分支)我们有参数表示 $x=\cosh t$, $y=\sinh t$. 对于由双曲线和参数值为 0 和 t 的点引出的射线所包围的二倍面积，我们确有

$$2A=\int_0^t(x\dot{y}-y\dot{x})d\tau=\int_0^t(\cosh^2\tau-\sinh^2\tau)d\tau$$

$$= \int_0^t d\tau = t.$$

(在射线上的积分值为零.)

I. 质量中心和曲线的矩

我们现在讨论力学中出现的一些概念. 我们考虑平面上 n 个质点组成的系统，它们有质量 $m_1, m_2, \cdots, m_n$，且坐标分别为 $y_1, y_2, \cdots, y_n$. 那么，我们称

$$T = \sum_{\nu=1}^{n} m_\nu y_\nu = m_1 y_1 + m_2 y_2 + \cdots + m_n y_n$$

为这质点系关于 x 轴的矩 (moment). 表示式 $\eta = T/M$ 定义了在 x 轴之上质点系的质量中心的高，即它的纵坐标，其中 M 表示系统的总质量 $m_1 + m_2 + \cdots + m_n$. 质量中心的高恰好是 $y_1, y_2, \cdots, y_n$ 的加权平均值，权因子是 $m_1, m_2, \cdots, m_n$ (见第 151 页). 因此，η 为质量的平均高度. 类似地，我们定义关于 y 轴的矩和质量中心的横坐标.

我们现在可以很容易地把这些矩的定义推广到质量均匀分布的曲线上. 因此，可以定义这样曲线的质量中心的坐标 ξ 和 η. (沿曲线密度为常数(譬如说 μ)的假定不是必要的，任何连续的分布都可以一样好地讨论.)

我们使用力学上典型的过程，从质点为有限个 n 的一个系统着手，然后令当 $n \to \infty$ 时取极限. 为此，我们引入弧长 s 作为曲线上的参数，并且用 $(n-1)$ 个分点把曲线分成长为 $\Delta s_1, \Delta s_2, \cdots, \Delta s_n$ 的弧. 我们把每一弧 Δs_i 的质量 $\mu \Delta s_i$ 假想地集中在弧上一任意点，譬如说在纵坐标为 y_i 的那一点.

根据定义，这个质点系关于 x 轴的矩为

$$T = \mu \sum y_i \Delta s_i.$$

如果量 Δs_i 之中最大的趋于零，那么这个和就趋于由积分

$$T = \mu \int_{s_0}^{s_1} y ds = \mu \int_{x_0}^{x_1} y \sqrt{1 + y'^2} dx \tag{23}$$

给出的极限，因此，我们很自然地把(23)作为曲线关于 x 轴的矩的定义. 因为曲线的全部质量等于它的长度乘以 μ,

$$\mu\int_{s_0}^{s_1} ds=\mu(s_1-s_0),$$

所以我们立即得到下面的曲线质量中心的坐标公式:

$$\xi=\frac{\int_{s_0}^{s_1} x\,ds}{s_1-s_0},\qquad \eta=\frac{\int_{s_0}^{s_1} y\,ds}{s_1-s_0}. \tag{24}$$

这些关系式实际上是曲线的矩和质量中心的定义；但是它们是有限个质点的比较简单情况的直接推广，从而我们很自然地期望——实际情况正是这样——在力学上，任何涉及质点系的质量中心和矩的关系式对于沿着曲线有连续的质量分布的情况也是正确的.

m. 旋转曲面的面积和体积

古鲁金定律

如果我们把曲线 $y=f(x)$，其中 $f(x)\geqslant 0$，绕 x 轴旋转，那么曲线就描绘出所谓的旋转曲面. 假定曲面的横坐标在 x_0 和 $x_1>x_0$ 之间，则它的面积可由类似于前面的讨论而得到. 因为如果我们用内接多边形代替曲线，我们就得到若干个细截锥构成的图形代替曲面. 由直观的启发，我们可以把旋转曲面的面积定义为当内接多边形最长边的长度趋于零时，这些锥面面积的极限.

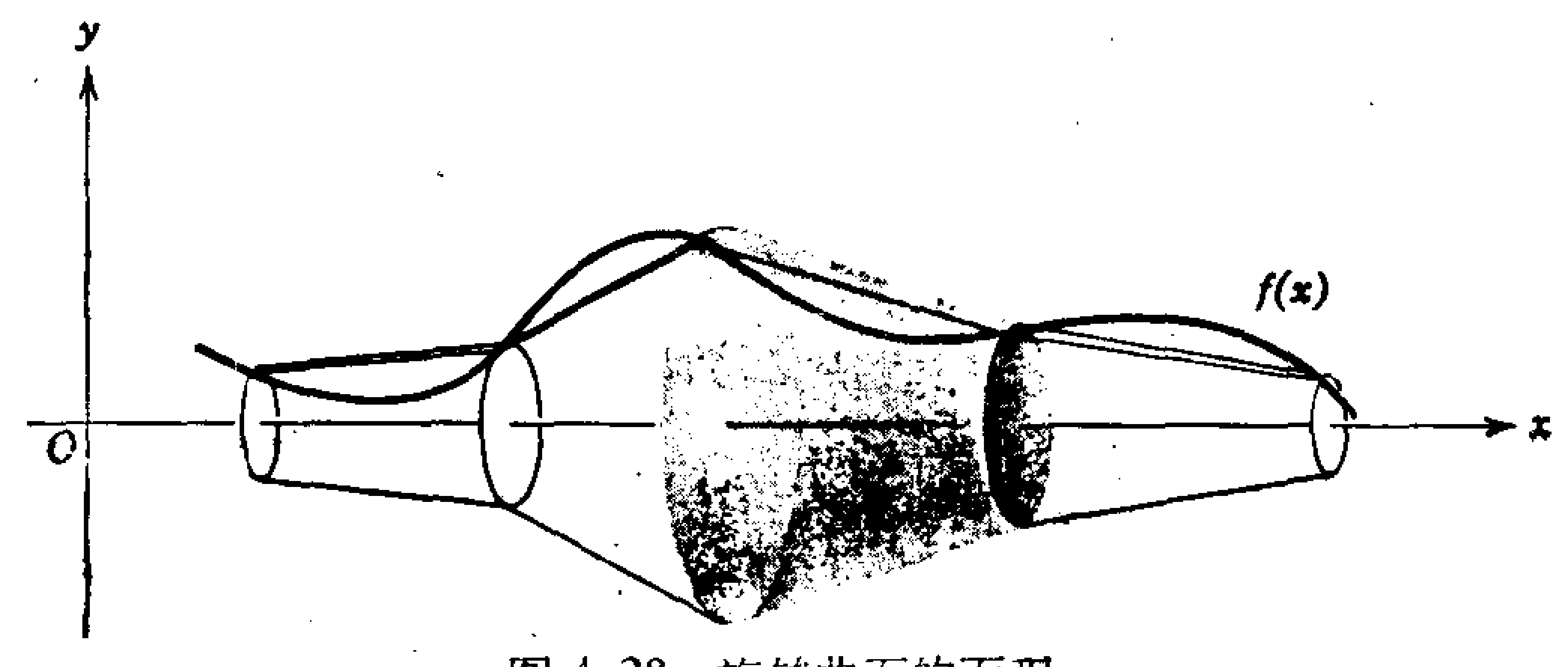


图 4.28 旋转曲面的面积

由初等几何我们知道，每一个截锥的面积等于斜母线的边长乘以平均半径的圆截面的周长(图 4.28). 如果我们把这些式子相加然后取极限，我们就得到面积的表达式

$$A = 2\pi\int_{s_0}^{s_1} y ds = 2\pi \int_{x_0}^{x_1} y \sqrt{1 + y'^2}dx = 2\pi\eta(s_1 - s_0). \quad (25)$$

用语言来表达，这个结果说明旋转曲面的面积等于母线的长度乘以质量中心走过的距离(古鲁金 (Guldin) 定律).

用同样的方法，我们可以发现，由旋转曲面和两端的平面 $x = x_0$ 和 $x = x_1 > x_0$ 所包围的内部体积为

$$V = \pi \int_{x_0}^{x_1} y^2 dx. \quad (26)$$

这个公式可以由下面直观的启发而得到，即所求的体积是早先提到的由截锥组成的图形体积的极限. 剩下的证明留给读者.

n. 惯性矩

在研究物体的旋转时，某些称为惯性矩的量起着重要的作用. 这里，我们简单地叙述这些表达式.

我们假定质点 m 在和 x 轴有距离 y 的地方，以角速度 ω (即在单位时间内旋转了角 ω) 绕该轴匀速旋转. 质点的动能表示为质量和速度平方的乘积的二分之一，等于

$$\frac{m}{2}(y\omega)^2.$$

我们把 $\frac{\omega^2}{2}$ 的系数 my^2 称为质点关于 x 轴的惯性矩.

类似地，如果我们有质量为 $m_1, m_2, \cdots, m_n$ 且纵坐标为 $y_1, y_2, \cdots, y_n$ 的 n 个质点，那么我们称表达式

$$T = \sum_i m_i y_i^2$$

为该质点系关于 x 轴的惯性矩. 惯性矩是属于质点系本身的一个量，与它的运动状态无关. 它的重要性在于，在系统围绕一个轴作保持每一对质点的距离不改变的刚体转动时，动能等于关于那个

轴的惯性矩乘以角速度平方的二分之一．因此，在围绕轴旋转时，关于该轴的惯性矩起了和质量在直线运动中同样的作用．

令 $y=f(x)$ 为位于横坐标 x_0 和 $x_1>x_2$ 之间的一条任意的曲线．沿着曲线质量以单位密度均匀分布．为了定义这条曲线的惯性矩，我们的作法完全象前节一样，得到关于 x 轴的惯性矩的表达式为

$$T_x=\int_{s_0}^{s_1} y^2 ds=\int_{x_0}^{x_1} y^2\sqrt{1+y'^2}dx. \tag{27}$$

关于 y 轴的惯性矩，我们相应地有

$$T_y=\int_{s_0}^{s_1} x^2 ds=\int_{x_0}^{x_1} x^2\sqrt{1+y'^2}dx. \tag{28}$$

4.2 例

我们从许多种平面曲线中选择几个典型的例子，来说明我们讨论过的概念．

a．普通摆线

根据方程 $x=a(t-\sin t)$，$y=a(1-\cos t)$（参见第353页(1)），有 $\dot{x}=a(1-\cos t)$，$\dot{y}=a\sin t$．我们得到弧长为

$$s=\int_0^\alpha\sqrt{\dot{x}^2+\dot{y}^2}dt=\int_0^\alpha\sqrt{2a^2(1-\cos t)}\,dt.$$

因为 $1-\cos t=2\sin^2\frac{t}{2}$，所以被积函数等于 $2a\sin\frac{t}{2}$，因而，对于 $0\leqslant\alpha\leqslant 2\pi$，有

$$s=2a\int_0^\alpha\sin\left(\frac{t}{2}\right)dt=-4a\cos\frac{t}{2}\Big|_0^\alpha$$

$$=4a\left(1-\cos\frac{\alpha}{2}\right)=8a\sin^2\frac{\alpha}{4}.$$

特别，如果我们考虑相继的两个尖点之间的弧长，那么我们必须令 $\alpha=2\pi$，因为参数值区间 $0\leqslant t\leqslant 2\pi$ 相应于滚动的圆的一次回转．因此，我们得到弧长的值为 $8a$；即相继两尖点之间的摆线的

弧长等于滚动圆直径的四倍.

类似地,我们计算以摆线的弧和 x 轴为边界的区域的面积:

$$I = \int_0^{2\pi} y\dot{x}dt = a^2\int_0^{2\pi} (1 - \cos t)^2 dt$$

$$= a^2\int_0^{2\pi} (1 - 2\cos t + \cos^2 t)dt$$

$$= a^2\left(t - 2\sin t + \frac{t}{2} + \frac{\sin 2t}{4}\right)\Big|_0^{2\pi} = 3a^2\pi.$$

所以这个面积是滚动圆面积的三倍.

对于曲率半径 $|\rho| = \frac{1}{|K|}$, 根据第 380 页(15)式我们有

$$\rho = \frac{(\dot{x}^2 + \dot{y}^2)^{\frac{3}{2}}}{\ddot{y}\dot{x} - \dot{y}\ddot{x}} = -2a\sqrt{2(1 - \cos t)} = -4a\left|\sin\frac{t}{2}\right|;$$

在点 $t = 0$, $t = \pm 2\pi$, $\cdots$, 上式为零. 这些点实际上是尖点. 在尖点上,摆线与 x 轴成直角.

根据第 399 页公式(25), 摆线的弧绕 x 轴旋转而形成的回转曲面的面积为

$$A = 2\pi\int_0^{8a} yds = 2\pi\int_0^{2\pi} a(1 - \cos t)\cdot 2a\sin\frac{t}{2}dt$$

$$= 8a^2\pi\int_0^{2\pi}\sin^3\frac{t}{2}dt = 16a^2\pi\int_0^{\pi}\sin^3 udu$$

$$= 16a^2\pi\int_0^{\pi}(1 - \cos^2 u)\sin udu.$$

作替换 $\cos u = v$ 可求出最后一个积分的值,我们得到

$$A = 16a^2\pi\left(-\cos u + \frac{1}{3}\cos^3 u\right)\Big|_0^{\pi} = \frac{64a^2\pi}{3}.$$

作为练习, 读者可以自己计算摆线在 x 轴上的质量中心的高 η 和惯性矩 T_x. 结果为

$$\eta = \frac{4}{3}a = \frac{A}{2\pi s}, \qquad T_x = \frac{256}{15}a^3.$$

b. 悬链线

悬链线[1]是由方程 $y = \cosh x$ 定义的曲线. 悬链线在横坐标 $x = a$ 和 $x = b$ 之间的长是

$$s = \int_a^b \sqrt{1 + \sinh^2 x}\, dx = \int_a^b \cosh x dx = \sinh b - \sinh a.$$

悬链线绕 x 轴旋转而产生的回转曲面的面积，即所谓的悬链曲面的面积为

$$A = 2\pi \int_a^b \cosh^2 x dx = 2\pi \int_a^b \frac{1 + \cosh 2x}{2} dx$$

$$= \pi\left(b - a + \frac{1}{2}\sinh 2b - \frac{1}{2}\sinh 2a\right).$$

由此我们进一步得到从 a 到 b 的弧的质量中心的高:

$$\eta = \frac{A}{2\pi s} = \frac{b - a + \frac{1}{2}\sinh 2b - \frac{1}{2}\sinh 2a}{2(\sinh b - \sinh a)}.$$

最后，曲率为

$$\kappa = \frac{y''}{(1 + y'^2)^{\frac{3}{2}}} = \frac{\cosh x}{\cosh^3 x} = \frac{1}{\cosh^2 x}.$$

c. 椭圆和双纽线

这两个曲线的弧长不能化为初等函数，而是属于第 319 页提到的椭圆积分.

对于椭圆 $y = (b/a)\sqrt{a^2 - x^2}$，我们有

$$s = \frac{1}{a}\int \sqrt{\frac{a^4 - (a^2 - b^2)x^2}{a^2 - x^2}}\, dx = a\int \sqrt{\frac{1 - \eta^2\xi^2}{1 - \xi^2}}\, d\xi,$$

---

1) 该名称来自这样一个事实：悬挂在两个端点上的一个链就是这个曲线的形状. 十分奇怪的是同一条曲线产生在很不同的物理应用中. 让我们看空间中的两个圆所限制的肥皂膜罢. 假定两个圆在互相平行的平面上并且通过两圆心的直线与这些平面垂直. 那么，这个肥皂膜的形状和悬链线绕 X 轴旋转而产生的回转曲面的形状完全一样.

其中我们置 $\frac{x}{a}=\xi$，$1-\frac{b^2}{a^2}=\eta^2$. 用代换 $\xi=\sin\phi$，这个积分就可表为

$$s=a\int\sqrt{1-\eta^2\sin^2\phi}\,d\phi.$$

这里，为得到椭圆的半周长，我们必须令 x 通过从 $-a$ 到 a 的区间，这个区间相应于区间

$$-1\leqslant\xi\leqslant+1\quad\text{或}\quad-\frac{\pi}{2}\leqslant\phi\leqslant+\frac{\pi}{2}.$$

对于双纽线，它在极坐标 r，t 下的方程为 $r^2=2a^2\cos 2t$，类似地有

$$s=\int\sqrt{r^2+\dot{r}^2}dt=\int\sqrt{2a^2\cos 2t+2a^2\frac{\sin^2 2t}{\cos 2t}}\,dt$$

$$=a\sqrt{2}\int\frac{dt}{\sqrt{\cos 2t}}=a\sqrt{2}\int\frac{dt}{\sqrt{1-2\sin^2 t}}.$$

如果在最后一个积分中引入 $u=\operatorname{tg} t$ 作为自变量，我们就有

$$\sin^2 t=\frac{u^2}{1+u^2},\quad dt=\frac{du}{1+u^2},$$

因此

$$s=a\sqrt{2}\int\frac{du}{\sqrt{1-u^4}}.$$

在双纽线的一个完全回路中，u 是由 -1 变到 $+1$，因此弧长等于

$$a\sqrt{2}\int_{-1}^{+1}\frac{du}{\sqrt{1-u^4}},$$

这个特殊的椭圆积分在高斯（Gauss）的研究中起了很大的作用.

4.3 二维向量

在讨论平面曲线与几何、力学和物理中其他许多论题时，向量概念已成为方便的、几乎不可缺少的工具，本章我们将发展和应用二维向量的概念，对于高维的推广放到第二卷叙述.

直观的解释

很多数学和物理的对象，在一个给定的尺度下可用一个单个的数完全地表示它,我们称它为标量（scalar）. 诸如角度、长度、面积、时间、质量和温度都是标量的例子. 但是还有另外一些对象,不可能只用一个标量表示,例如,三角形的形状、空间中点的位置、质点运动的方向或加速度以及物体的张力. 需要几个数才能确定每一个这样的对象. 逐渐地,数学概念超出实数连续统而发展起来,使我们能够用单个的记号来表示这种对象.[1]平面上的向量(vector)可以用两个信息项来描述:即长度和方向.例如,两点的相对位置、质点的速度和加速度与作用在一个质点上的力都属于这一类[2].

几何上或直观地，向量在本质上是用它的长(或大小)和它的方向来描述的平面(或空间)上的有向线段. 通常，向量用给定的长和指向给定的方向的箭头符号表示. 除非明确加以限制，向量是“自由的”，也就是说在向量的定义中并不规定有向线段起点的位置.

虽然有许多物理概念,例如速度、加速度和力是向量在应用中的基本例子,但是我们将用平移或平行位移从几何上来定义向量.

向量分析是从给有向线段或平行位移命名为“向量”开始的,但是其决定性的意义不是引入了一个统一的名词,而是这些对象,即向量(类似的,如复数),可借助一套法则而互相结合,或与标量互相结合,这套法则称为向量代数或向量分析,它们在各种应用中有着自然的解释,例如,两个速度的叠加,或位移反抗一个力所作的功. 直观借助的向量的语言可使许多的数学和物理关系表达得简洁而清楚.

a. 用平移定义向量. 记号

平面上最简单的变换是平移，即平行位移. 平移是把任意一

---

1) 当然,复数 $a+bi=z$ 是表示实数对 a, b 的这样的符号,有时用复数而不用向量的确方便些.

2) 有时向量也是不够的, 譬如描述张力或空间曲率就要使用更一般的称为张量的概念.

点 $P=(x, y)$ 变到或者映射到点 $P'=(x', y')$，P' 的坐标是

$$x'=x+a, \quad y'=y+b,$$

其中 a 和 b 为常数. 平移由常数 a 和 b 完全决定. 我们把 a 和 b 称为平移的分量，(component). 我们将用"向量"一词作为平移的另一个名称. 使用黑体字表示向量或平移. 我们把分量为 a, b 的向量记为 $\mathbf{R}=(a, b)$ (图 4.29).

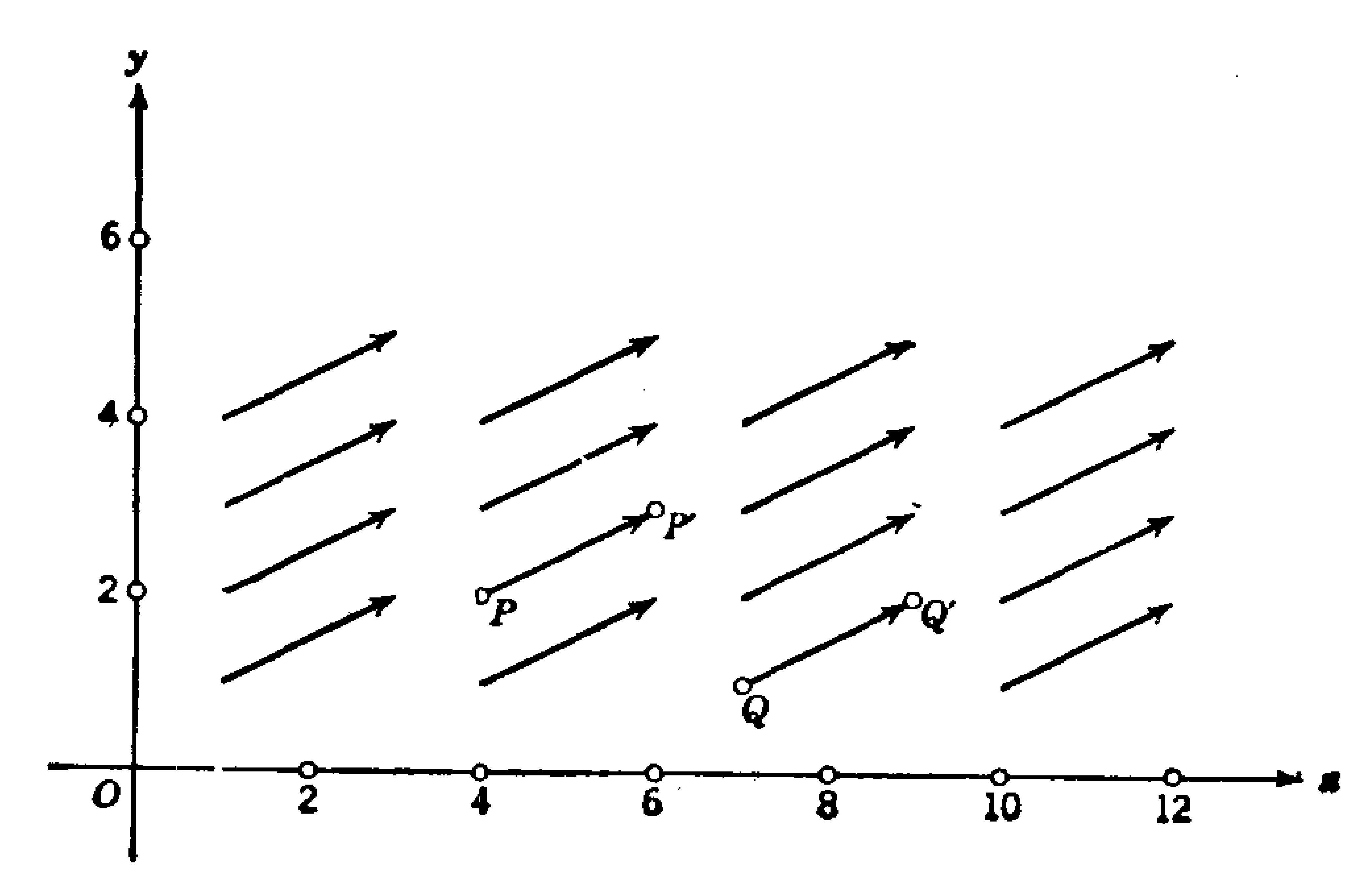


图 4.29 对应于向量 $\mathbf{R}=\overrightarrow{PP'}=\overrightarrow{QQ'}=(2, 1)$ 的平移 $x'=x+2$, $y'=y+1$

向量 $\mathbf{R}$ 的分量由一对相应的点 $P=(x, y)$ 和 $P'=(x', y')$ 决定，因为

$$a=x'-x, \quad b=y'-y.$$

显然，对于任意点 P 和 P' 总能找到使 P 变到 P' 的平移 $\mathbf{R}$. 我们把它表示为向量 $\mathbf{R}=\overrightarrow{PP'}$. 因此，任意有序点对 $P=(x, y)$, $P'=(x', y')$, 即任意有向线段决定向量 $\mathbf{R}=\overrightarrow{PP'}=(x'-x, y'-y)$. 我们看到，第二对点 $Q=(\xi, \eta)$, $Q'=(\xi', \eta')$ 当 $\xi'-\xi=x'-x$, $\eta'-\eta=y'-y$ 时定义了同一个向量；因此，同样的平移 $\mathbf{R}$ 使 P 变到 P'，Q 变到 Q'. 向量 $\mathbf{R}$ 由两个数 (即分量) 决定，就像平面上的点由两个坐标决定一样. 基本区别是在几何上用一个点对表示向量，在表示式 $\mathbf{R}=\overrightarrow{PP'}$ 中，我们把 P 称为起点，把 P' 称为终点. 对于给定的 $\mathbf{R}$，两点之一，譬如说起点 $P=$

(x, y) 可以任意选定；而后终点 $P'=(x', y')$ 就由关系 $x'=x+a$, $y'=y+b$ 唯一地确定. 把起点和终点交换位置就得到反向量 $\overrightarrow{P'P}=(-a,-b)$.

如果把起点选在原点 $O=(0, 0)$, 那么我们可以取 $\mathbf{R}=\overrightarrow{OQ}$ 而把向量 $\mathbf{R}$ 与每一个点 $Q=(x, y)$ 唯一地联系起来. 有固定的起点 O 的向量称为 Q 的位置向量. Q 的位置向量的分量就是 Q 的坐标 x, y.

分量为 $a=0$, $b=0$ 的向量 $\mathbf{R}$ 叫做零向量，记为 $\mathbf{O}$. 它相当于一个把每一个点保持固定的平移:

$$\mathbf{O}=(0, 0)=\overrightarrow{PP}.$$

两点 $P=(x, y)$, $P'=(x', y')$ 的距离 r 只依赖于向量 $\mathbf{R}=(a, b)=\overrightarrow{PP'}$, 因为

$$r=\sqrt{(x'-x)^2+(y'-y)^2}=\sqrt{a^2+b^2}.$$

我们把 r 称为向量 $\mathbf{R}$ 的长，并且记为 $r=|\mathbf{R}|$. 除非 $\mathbf{R}=0$, $\mathbf{R}$ 的长总是一个正数(见图 4.30).

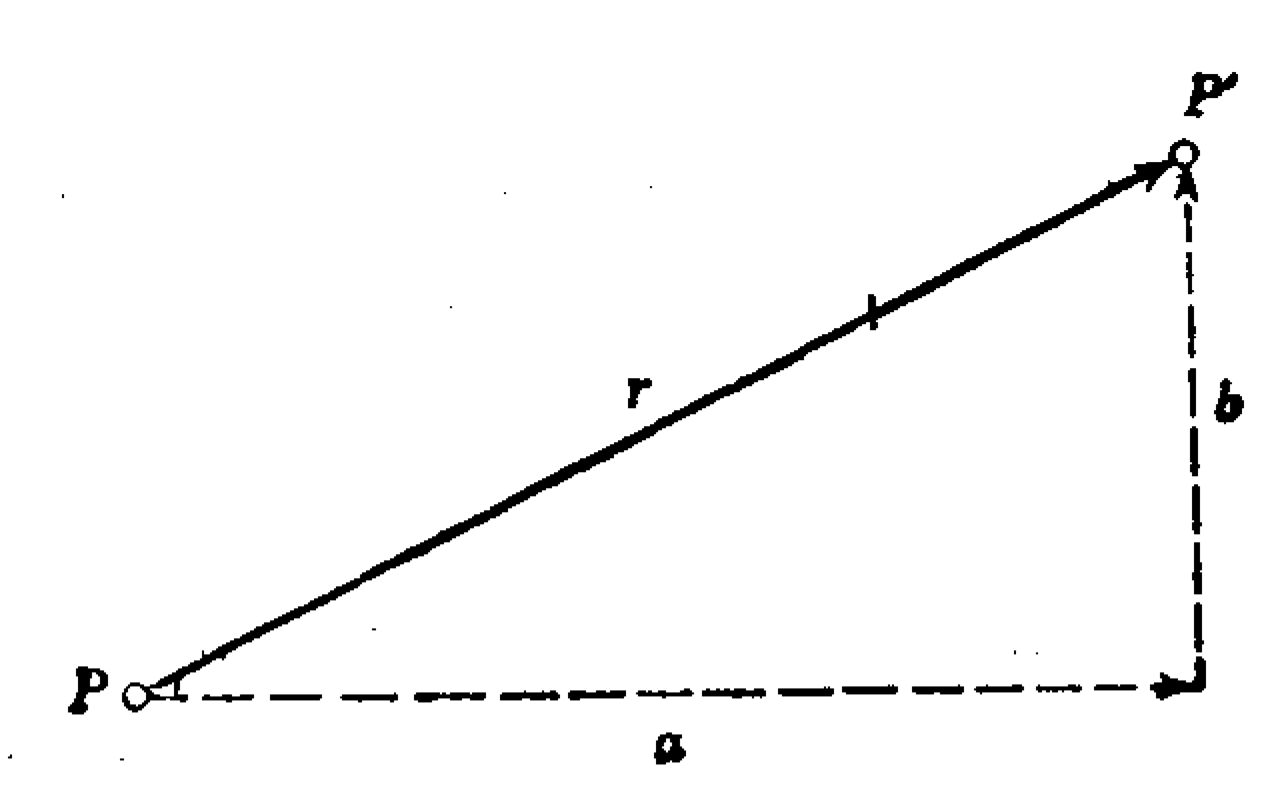


图 4.30 向量 $\mathbf{R}=\overrightarrow{PP'}$ 的分量 a, b 和长 r

我们把向量 $\mathbf{R}=(a, b)$ 与数或标量 λ 的乘积定义为向量

$$\mathbf{R}^*=\lambda\mathbf{R}=(\lambda a, \lambda b).$$

如 $\lambda=-1$, 我们有和 $\mathbf{R}$ 相反的向量 $\mathbf{R}^*=(-a,-b)$(图 4.31).

如果 $\mathbf{R}=\overrightarrow{PP'}=(a, b)$, $P=(x, y)$, $P'=(x', y')$, 那么我们可以把 $\mathbf{R}^*=\lambda\mathbf{R}$ 表为 $\overrightarrow{PP''}$, 其中 $P''=(x'', y'')=(x+\lambda a, y+\lambda b)$ (见图 4.32). 如 $a=b=0$, 我们当然有 $P''=P'=P$. 如 a 和 b 不都是零，则点 $P''=(x'', y'')=(x+\lambda a, y+\lambda b)$,

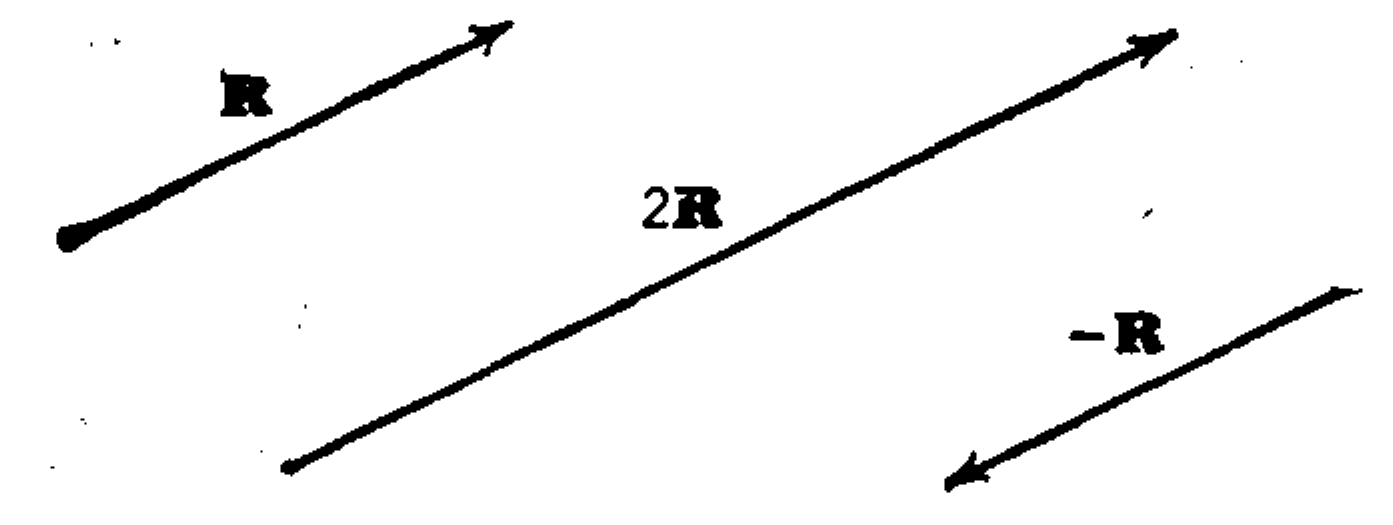


图 4.31　向量 R 与标量相乘

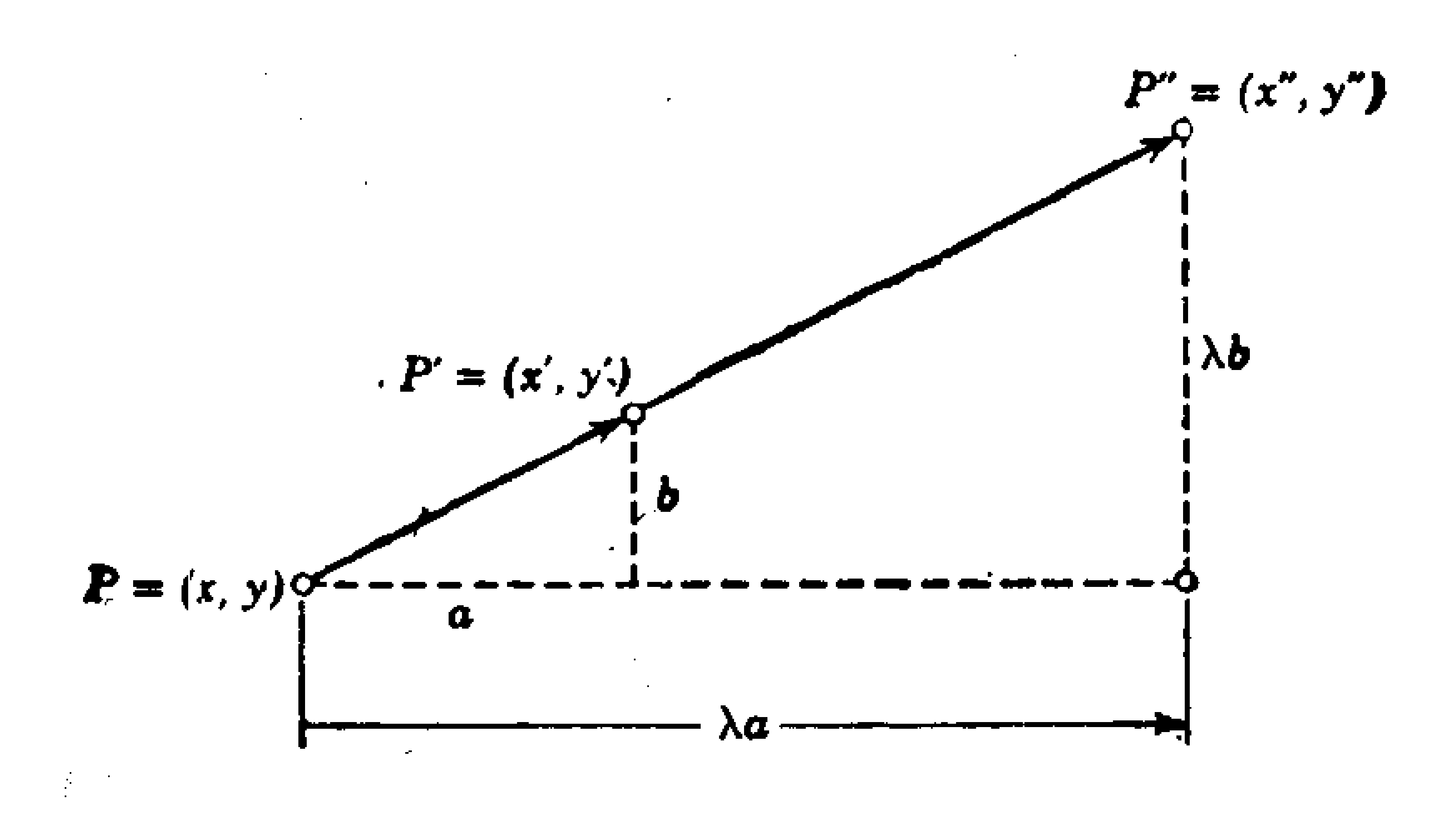


图 4.32　在 $\lambda=\frac{8}{3}$ 时的向量关系 $\mathbf{R}^*=\overrightarrow{PP''}=\lambda\overrightarrow{PP'}$

当 λ 取遍所有实数值时跑遍这整个直线

$$x''b-y''a=xb-ya.$$

如 $\lambda=0$，则 $P''=P$；如 $\lambda=1$，则 $P''=P'$. 因此 P'' 在过 P 和 P' 的直线上. 如 $\lambda>0$，那么点 P'' 和 P' 在 P 的同一边；如 $\lambda<0$，P'' 和 P' 在 P 的相反的两边.

给定 $\mathbf{R}=(a,b)$ 和 $\mathbf{R}^*=(a^*,b^*)$. 如果 $\mathbf{R}^*=\lambda\mathbf{R}$ 并且 $\lambda>0$，那么就说 $\mathbf{R}$ 和 $\mathbf{R}^*$ 有相同的方向；如 $\lambda<0$，就说 $\mathbf{R}$ 和 $\mathbf{R}^*$ 有相反的方向. 如果 $\mathbf{R}=0$，那么同样 $\mathbf{R}^*=0$. 如果 $\mathbf{R}\neq0$，那么 $\mathbf{R}^*$ 和 $\mathbf{R}$ 有相同方向的充分必要条件是

$$\frac{a}{\sqrt{a^2+b^2}}=\frac{a^*}{\sqrt{a^{*2}+b^{*2}}},\qquad \frac{b}{\sqrt{a^2+b^2}}=\frac{b^*}{\sqrt{a^{*2}+b^{*2}}}.$$

我们把决定向量 $\mathbf{R}$ 的方向的量

$$\xi=\frac{a}{\sqrt{a^2+b^2}}=\frac{a}{|\mathbf{R}|}=\frac{a}{r},\quad \eta=\frac{b}{\sqrt{a^2+b^2}}=\frac{b}{|\mathbf{R}|}=\frac{b}{r}$$

叫做 $\mathbf{R}$ 的方向余弦；当然它们对于 $\mathbf{R}=0$ 没有定义. 因为 ξ^2+

$\eta^2=1$，所以我们总能找到角 α 和相应的角 $\beta=\frac{\pi}{2}-\alpha$，使得

$$\xi=\cos\alpha,\quad \eta=\sin\alpha=\cos\beta.$$

角 α 叫做 $\mathbf{R}$ 的方向角(图 4.33).除去 π 的偶数倍之外方向角是唯一地确定的. 对于 $\mathbf{R}=\overrightarrow{PP'}$，我们有

$$\cos\alpha=\frac{x'-x}{r},\qquad \sin\alpha=\frac{y'-y}{r}.$$

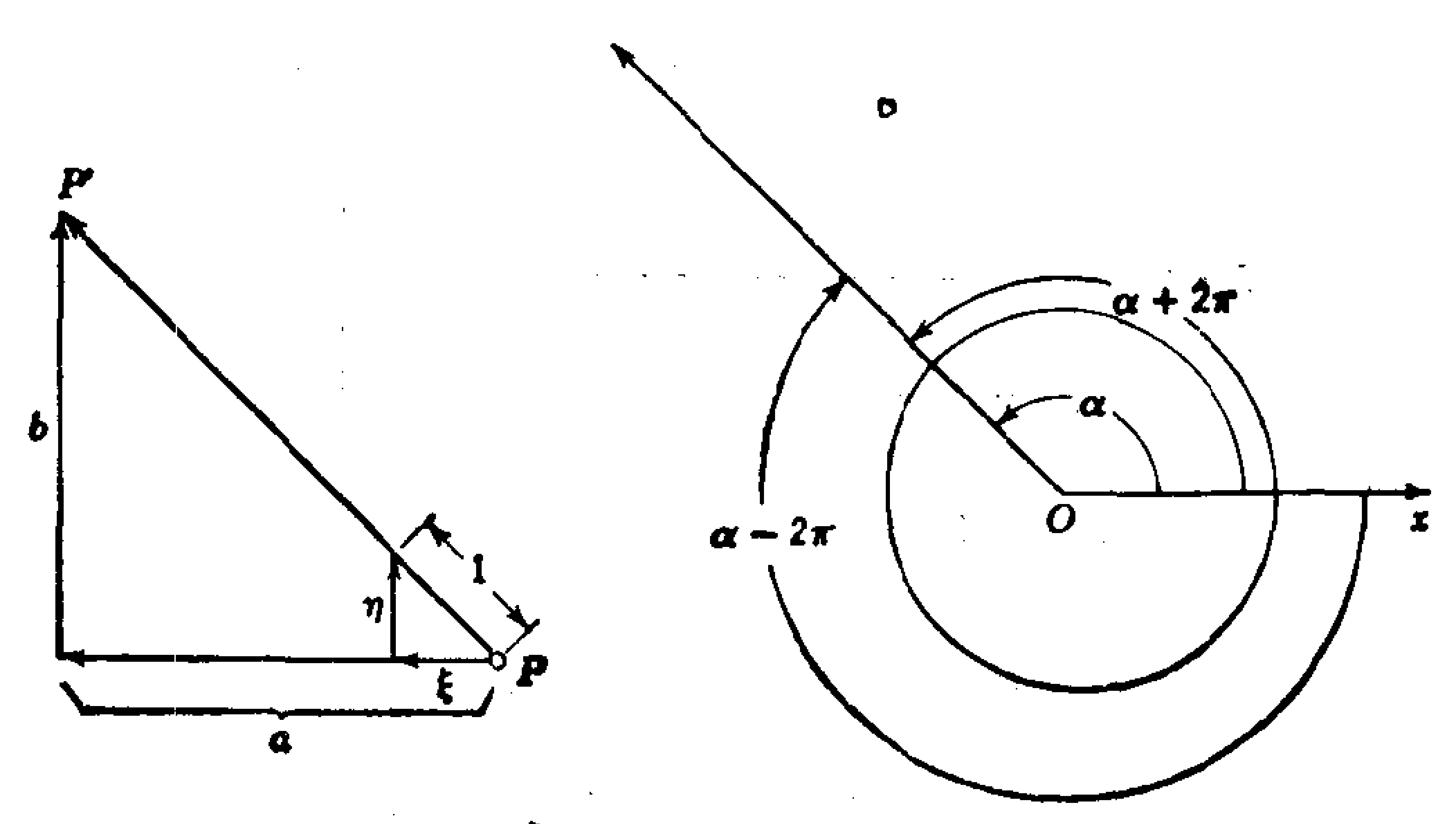


图 4.33　向量 $\overrightarrow{PP'}$ 的方向角和方向余弦 ξ,η

显然，α 为正 x 轴与 P 到 P' 的直线的夹角. 更确切地说，把正 x 轴绕原点旋转角 α(如反时针转，算正的；如顺时针转，算负的)该轴将给出由 P 到 P' 的方向.反向量 $-R=(-a,-b)$ 的方向余弦为 $-\xi,-\eta$,并且方向角与 α 相差 π 的奇数倍.如果向量 $\mathbf{R}=\overrightarrow{PP'}$ 的起点是原点，那么 $\mathbf{R}$ 的方向角 α 就是 P' 的极角 θ.

b. 向量的加法和乘法

向量的和

我们已经用平移，即平面上点的某些映射定义了向量. 存在一个完全一般的方法，通过相继地使用任意两个映射，将它们结合起来. 如果第一个映射把点 P 移到点 P'，第二个映射把点 P' 移到 P''，那么组合的映射是把 P 移到 P''. 在两个向量 $\mathbf{R}=(a,b)$ 和

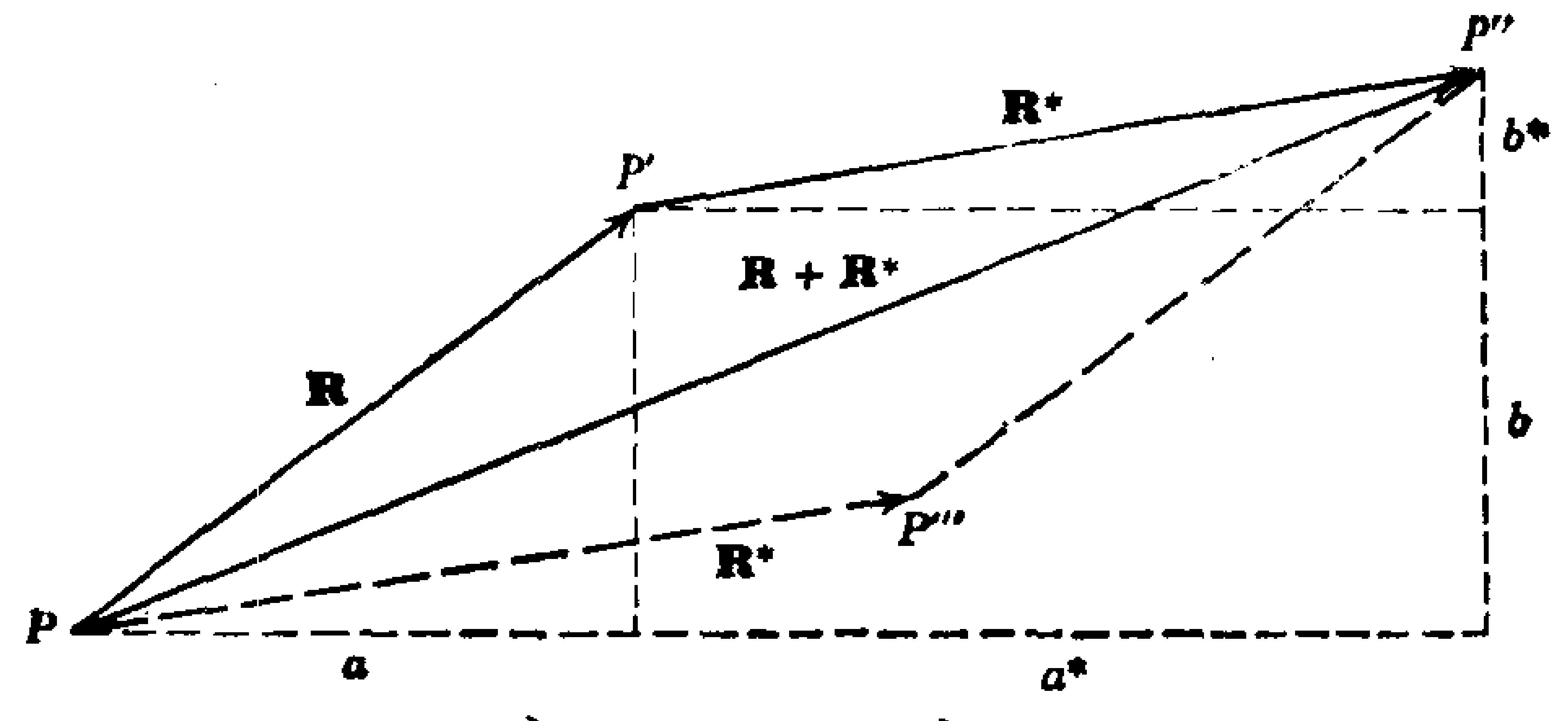


图 4.34 向量 $\overrightarrow{PP'}=(a,b)$ 和 $\overrightarrow{P'P''}=(a^*,b^*)$ 的加法

$\mathbf{R}^*=(a^*,b^*)$ 的情况下，向量 $\mathbf{R}$ 把点 $P=(x,y)$ 映射到点 $P'=(x+a,y+b)$，向量 $\mathbf{R}^*$ 把 P' 映射到点 $P''=(x+a+a^*,y+b+b^*)$. 从 P 到 P'' 产生的映射还是一个平移；我们称它为向量 $\mathbf{R}=\overrightarrow{PP'}$ 和 $\mathbf{R}^*=\overrightarrow{P'P''}$ 的和，记为 $\mathbf{R}+\mathbf{R}^*$（图 4.34）.[1]这个和向量的分量是 $a+a^*$ 和 $b+b^*$. 因此，两个向量的和定义为

$$\overrightarrow{PP'}+\overrightarrow{P'P''}=\overrightarrow{PP''},$$

或者使用分量描述，便是

$$(a,b)+(a^*,b^*)=(a+a^*,b+b^*).$$

如果 $\mathbf{R}^*$ 和 $\mathbf{R}$ 有相同起点，譬如说 $\mathbf{R}^*=\overrightarrow{PP'''}$，那么点 P,P''',P'' 和 P' 构成一个平行四边形的顶点. 过 P 的两个边表示向量 $\mathbf{R}$ 和 $\mathbf{R}^*$；过 P 点的对角线表示和 $\mathbf{R}+\mathbf{R}^*$（向量和的"平行四边形作图法"）.

向量和满足算术的交换律和结合律，因为向量的加法相当于对应的分量的加法（图 4.35）. 而且它们还满足两个向量之和乘以数 λ 和一个向量乘以两个数 λ,μ 之和的分配律：

$$\lambda(\mathbf{R}+\mathbf{R}^*)=\lambda\mathbf{R}+\lambda\mathbf{R}^*,\quad (\lambda+\mu)\mathbf{R}=\lambda\mathbf{R}+\mu\mathbf{R}\text{[2]}.$$

这些法则使我们可以用点 P 和 P' 的位置向量 $\overrightarrow{OP}$ 和 $\overrightarrow{OP'}$ 来表示向量 $\overrightarrow{PP'}$（图 4.36）：

1）"和"实际上是第 54 页定义的两个映射的符号乘积. 在这里加号更自然一些，因为它相应于分量的加法.

2）为了在方程中区别向量和数，在写乘积时，我们总是把数放在向量的前面，虽然可以定义 $\lambda\mathbf{R}=\mathbf{R}\lambda$，但我们不使用 $\mathbf{R}\lambda$ 的记法.

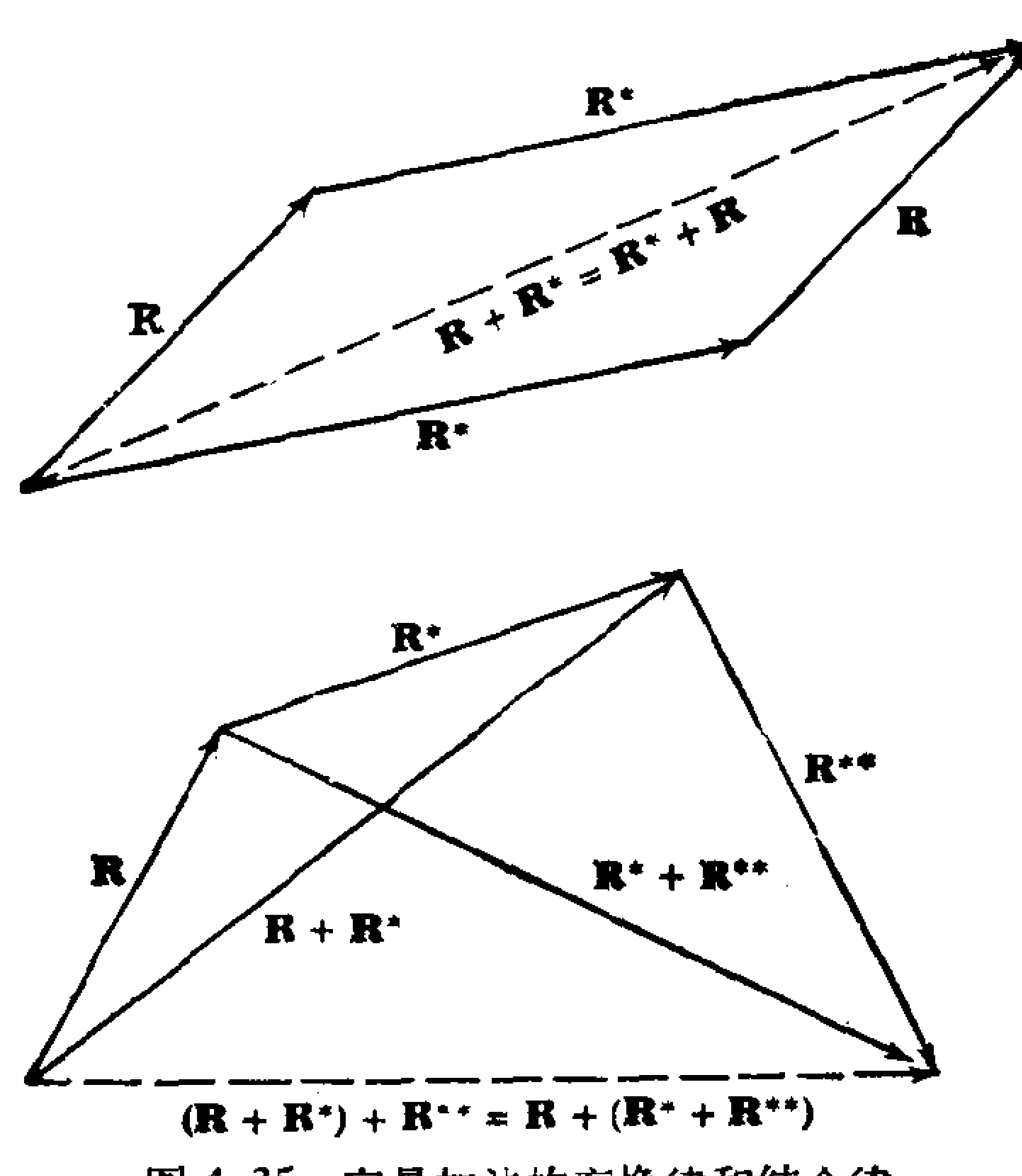


图 4.35　向量加法的交换律和结合律

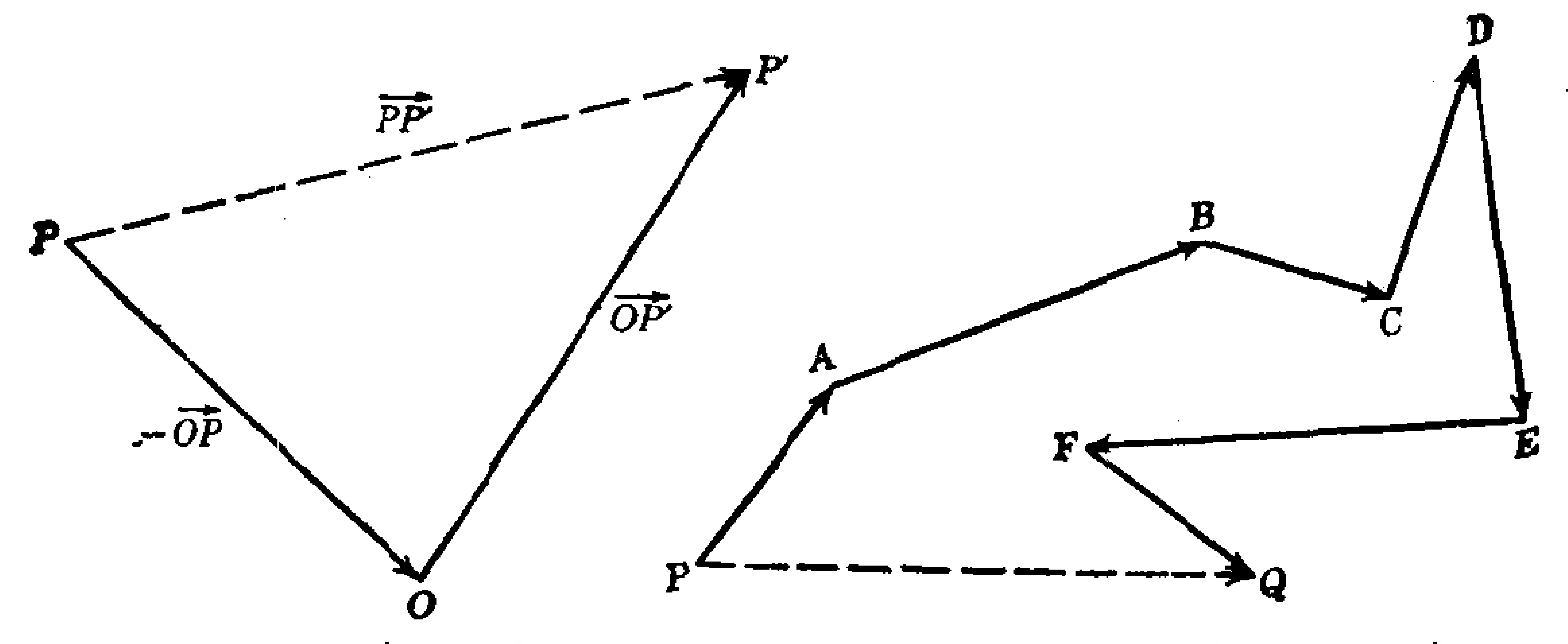


图 4.36　$\overrightarrow{PP'} = \overrightarrow{OP'} - \overrightarrow{OP}$　　图 4.37　$\overrightarrow{PQ} = \overrightarrow{PA} + \overrightarrow{AB} + \overrightarrow{BC} + \cdots + \overrightarrow{FQ}$

$$\overrightarrow{PP'} = \overrightarrow{PO} + \overrightarrow{OP'} = \overrightarrow{OP'} + \overrightarrow{PO} = \overrightarrow{OP'} - \overrightarrow{OP}.$$

重要的是认识到：一般地，如果我们从点 P 出发经过点 A，B，C,…,E, F 进行到 Q，那么向量 $\overrightarrow{PQ}$ 就是向量 $\overrightarrow{PA}$，$\overrightarrow{AB}$，$\overrightarrow{BC}$，…，$\overrightarrow{EF}$，$\overrightarrow{FG}$ 之和（图 4.37）.

向量之间的夹角

向量 $\mathbf{R}^* = (a^*, b^*)$ 和向量 $\mathbf{R} = (a, b)$ 的夹角 θ 定义为它们的方向角之差：$\theta = \alpha^* - \alpha$.（这里，假定 $\mathbf{R}$ 和 $\mathbf{R}^*$ 都不是零向

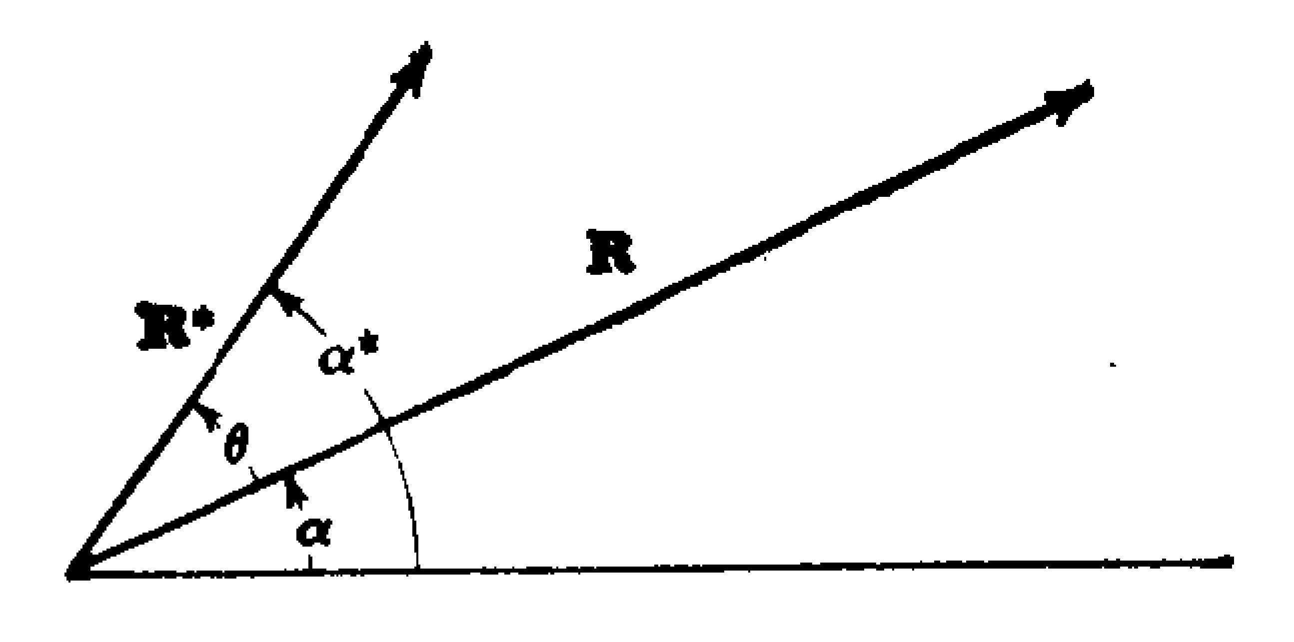


图 4.38　向量 $\mathbf{R}^*$ 与向量 $\mathbf{R}$ 构成角 θ

量.）除去可以相差 2π 的整数倍之外角 θ 仍是确定的（图 4.38）.旋转角 θ（θ 的正负号指示了旋转的方向）使 $\mathbf{R}$ 的方向变到 $\mathbf{R}^*$ 的方向. 唯一确定的量 $\cos\theta$ 和 $\sin\theta$ 可直接用 $\mathbf{R}$ 和 $\mathbf{R}^*$ 的方向余弦表示:

$$\cos\theta = \cos(\alpha^* - \alpha) = \cos\alpha\sin\alpha^* + \sin\alpha\sin\alpha^*$$
$$= \frac{aa^* + bb^*}{\sqrt{a^2 + b^2}\sqrt{a^{*2} + b^{*2}}},$$
$$\sin\theta = \sin(\alpha^* - \alpha) = \cos\alpha\sin\alpha^* - \sin\alpha\cos\alpha^*$$
$$= \frac{ab^* - a^*b}{\sqrt{a^2 + b^2}\sqrt{a^{*2} + b^{*2}}}.$$

每一式子的分母刚好是向量长的乘积 rr'. 我们引入在分子中出现的式子作为两个向量的“乘积”.

两个向量的内积和外积

我们定义向量 $\mathbf{R}=(a, b)$ 和 $\mathbf{R}^*=(a^*, b^*)$ 的“数量”积(又称“内”积或“点”积）为

$$\mathbf{R}\cdot\mathbf{R}^* = aa^* + bb^* = rr^*\cos\theta.$$

定义 $\mathbf{R}$ 和 $\mathbf{R}^*$ 的“外”积（或“叉”积）为

$$\mathbf{R}\times\mathbf{R}^* = ab^* - a^*b = rr^*\sin\theta.$$ 1)

由直接证明可知,**内积和外积满足分配律和结合律**:

$$\mathbf{R}\cdot(\mathbf{R}^* + \mathbf{R}^{**}) = \mathbf{R}\cdot\mathbf{R}^* + \mathbf{R}\cdot\mathbf{R}^{**},$$
$$\mathbf{R}\times(\mathbf{R}^* + \mathbf{R}^{**}) = \mathbf{R}\times\mathbf{R}^* + \mathbf{R}\times\mathbf{R}^{**},$$

1）按这里的定义，内积和外积实际上都是“标量”. 我们仅对于内积保留了数量积这个词,因为在三维空间中,与外积相类似的是向量.

$$\lambda(\mathbf{R}\cdot\mathbf{R}^*)=(\lambda\mathbf{R})\cdot\mathbf{R}^*=\mathbf{R}\cdot(\lambda\mathbf{R}^*),$$
$$\lambda(\mathbf{R}\times\mathbf{R}^*)=(\lambda\mathbf{R})\times\mathbf{R}^*=\mathbf{R}\times(\lambda\mathbf{R}^*).$$

对于内积，乘法的交换律也成立，

$$\mathbf{R}\cdot\mathbf{R}^*=\mathbf{R}^*\cdot\mathbf{R};$$

但是对于外积，如果交换因子，则方向相反：

$$\mathbf{R}\times\mathbf{R}^*=-\mathbf{R}^*\times\mathbf{R}.$$

令 $\mathbf{R}$ 和 $\mathbf{R}^*$ 有同一个起点，$\mathbf{R}=\overrightarrow{PQ}$，$\mathbf{R}^*=\overrightarrow{PQ^*}$. 我们可以把 $\mathbf{R}\cdot\mathbf{R}^*$ 解释为线段 PQ^* 在线段 PQ 上的射影 $r^*\cos\theta$ 和该线段长的乘积，外积 $\mathbf{R}\times\mathbf{R}^*$ 不过是有向三角 PQQ^* 的面积的两倍. 如果顶点 PQQ^* 是反时针的次序，面积取正号，反之取负号（图 4.39）.

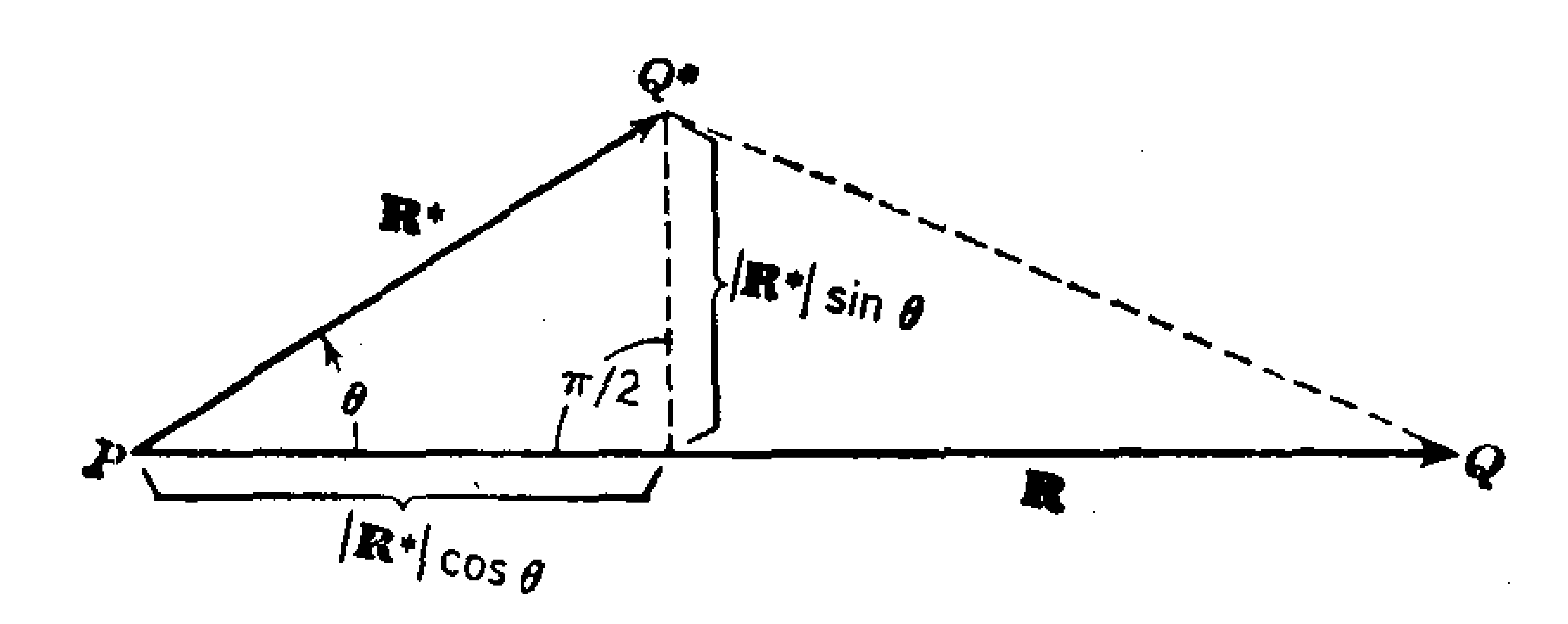


图 4.39　向量积 $\mathbf{R}\times\mathbf{R}^*=|\mathbf{R}||\mathbf{R}^*|\sin\theta$ 为三角形 PQQ^* 的面积的两倍

对任意向量 $\mathbf{R}=(a, b)$，

$$\mathbf{R}\cdot\mathbf{R}=a^2+b^2=|\mathbf{R}|^2$$

为向量长的平方. 因此，$\mathbf{R}\cdot\mathbf{R}$ 是正值，除非 $\mathbf{R}=0$. 另一方面，$\mathbf{R}\times\mathbf{R}$ 总为零. 两个非零向量互相正交的条件是 $\mathbf{R}\cdot\mathbf{R}^*=0$，而如果 $\mathbf{R}\times\mathbf{R}^*=0$，那么 $\mathbf{R}$ 和 $\mathbf{R}^*$ 是平行的（即方向相同或相反）.

直线的方程

使用向量记号，我们可以很容易写出过两点的直线方程和过一给定点有给定的方向的直线方程. 令 $P=(x, y)$，$P_0=(x_0, y_0)$ 和 $P_1=(x_1, y_1)$ 为三个点，并且 $P_0\neq P_1$. 如果 $\overrightarrow{P_0P}$ 和 $\overrightarrow{P_0P_1}$ 是平

行的，即

$$\overrightarrow{P_0P} \times \overrightarrow{P_0P_1} = 0,$$

那么 P 在过 P_0 和 P_1 的直线上．如果 $\mathbf{R} = \overrightarrow{OP}$，$\mathbf{R}_0 = \overrightarrow{OP_0}$ 和 $\mathbf{R}_1 = \overrightarrow{OP_1}$为三点的位置向量．那么条件就变为

$$(\mathbf{R} - \mathbf{R}_0) \times (\mathbf{R}_1 - \mathbf{R}_0) = 0$$

或

$$(\mathbf{R}_1 - \mathbf{R}_0) \times \mathbf{R} = \mathbf{R}_1 \times \mathbf{R}_0.$$

把点的坐标代替位置向量，我们就得到通常形式的直线方程（图4.40）：

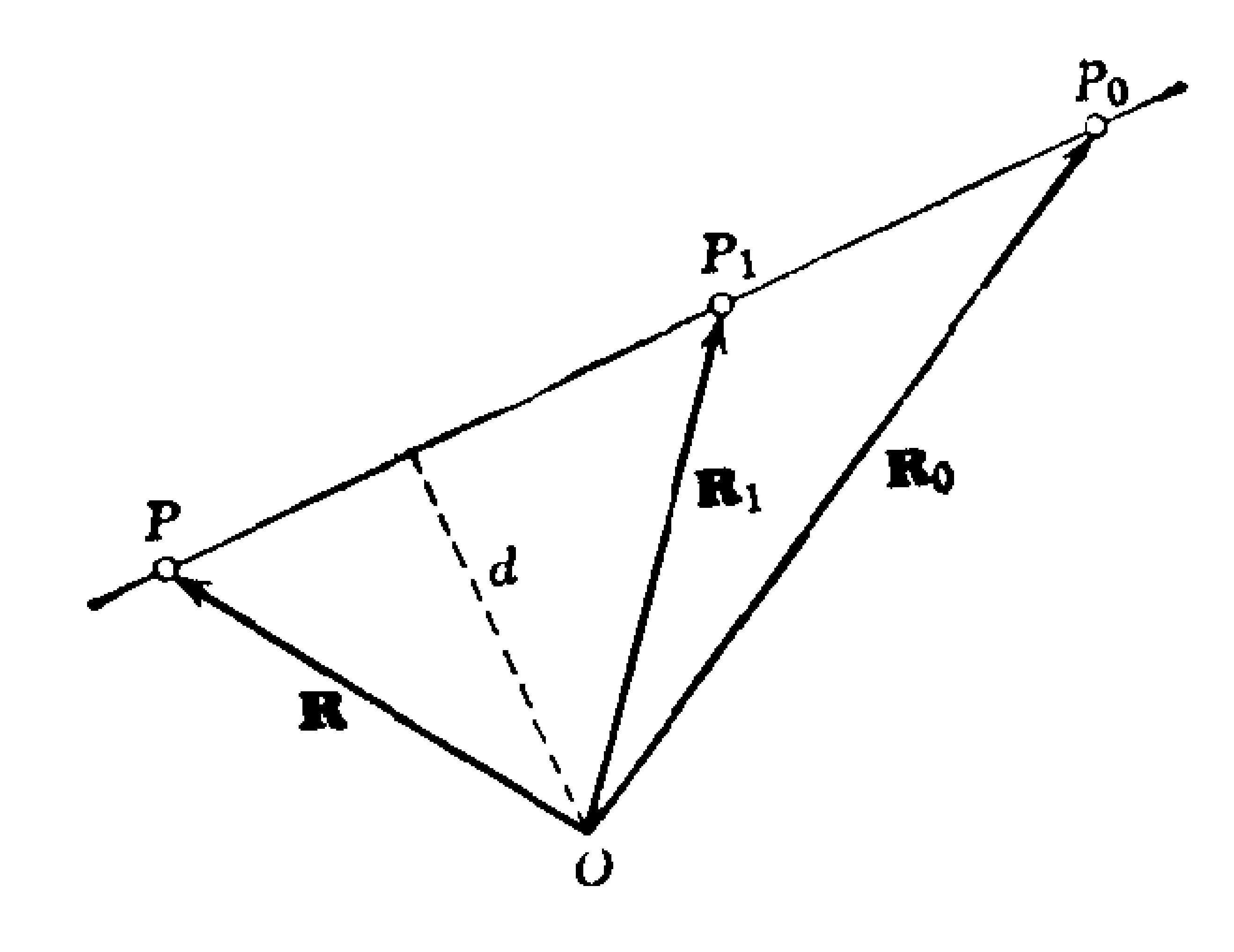


图 4.40　用向量符号表示的直线

$$(x_1 - x_0)y - (y_1 - y_0)x = x_1y_0 - y_1x_0.$$

如果不是指定直线上的两个点，我们也可以先定下一个点 P_0 来求与向量 $\mathbf{S} = (a, b)$ 平行的直线．显然，直线方程为

$$(\mathbf{R} - \mathbf{R}_0) \times \mathbf{S} = 0$$

即

$$(x - x_0)b - (y - y_0)a = 0.$$

如 $\mathbf{S} = \overrightarrow{P_0P_1}$，就得到前面的那个方程．

原点到直线的距离 d 也可以用向量符号表示．显然，d 乘以向量 $\overrightarrow{P_0P_1}$ 的长等于三角形 OP_0P_1 面积的两倍．因此

$$d = \frac{1}{|\overrightarrow{P_0P_1}|}(\overrightarrow{OP_0} \times \overrightarrow{OP_1}) = \frac{\mathbf{R}_0 \times \mathbf{R}_1}{|\mathbf{R}_1 - \mathbf{R}_0|}$$

$$= \frac{x_0y_1 - x_1y_0}{\sqrt{(x_1-x_0)^2+(y_1-y_0)^2}}.$$

这里，如果点 O，P_0，P_1 是反时针顺序的，d 就取正值。

坐标向量.

向量 $\mathbf{R}=(a, b)$ 通常可以表示为

$$\mathbf{R}=a\mathbf{i}+b\mathbf{j}, \tag{29}$$

这里 $\mathbf{i}$ 和 $\mathbf{j}$ 表示"坐标向量 (coordinate vectors)"

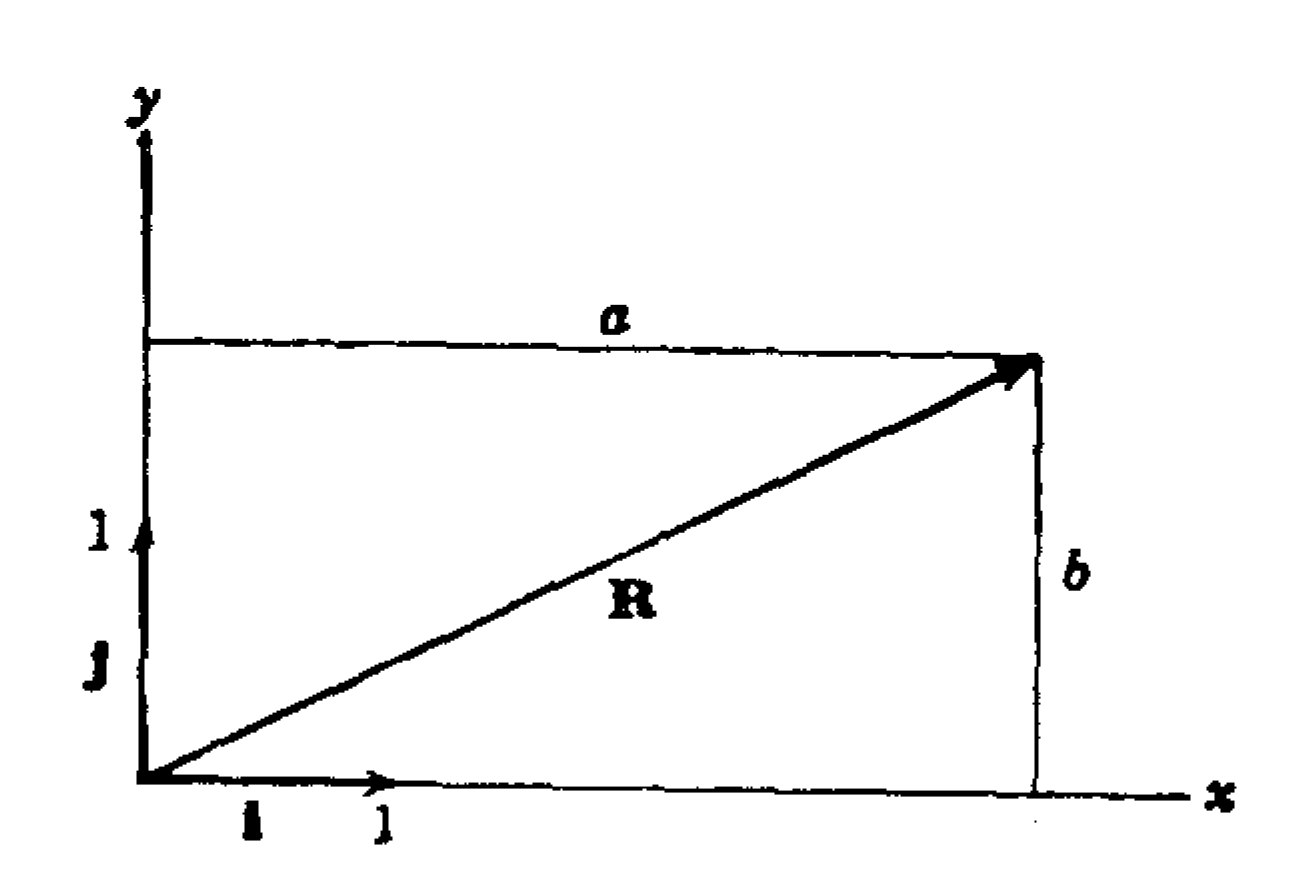
$$\mathbf{i}=(1, 0), \quad \mathbf{j}=(0, 1). \tag{30}$$

用这种方法可以把 $\mathbf{R}$ 分成两个分别平行于 x 轴和 y 轴方向的向量 $a\mathbf{i}$ 和 $b\mathbf{j}$. $\mathbf{R}$ 的分量 a 和 b 刚好是这两个向量的(有符号的)长.

在应用中，我们常常提出把向量 $\mathbf{R}$ 表成为具有两个给定的正交方向(即互相垂直)的向量的和. 为此目的，我们引入两个具有给定方向的单位向量 $\mathbf{I}$ 和 $\mathbf{J}$. 如果我们可以把 $\mathbf{R}$ 表为

$$\mathbf{R}=A\mathbf{I}+B\mathbf{J}, \tag{31}$$

这里 A，B 是两个数，那么就可以得到所要求的 $\mathbf{R}$ 的分解(图 4.40



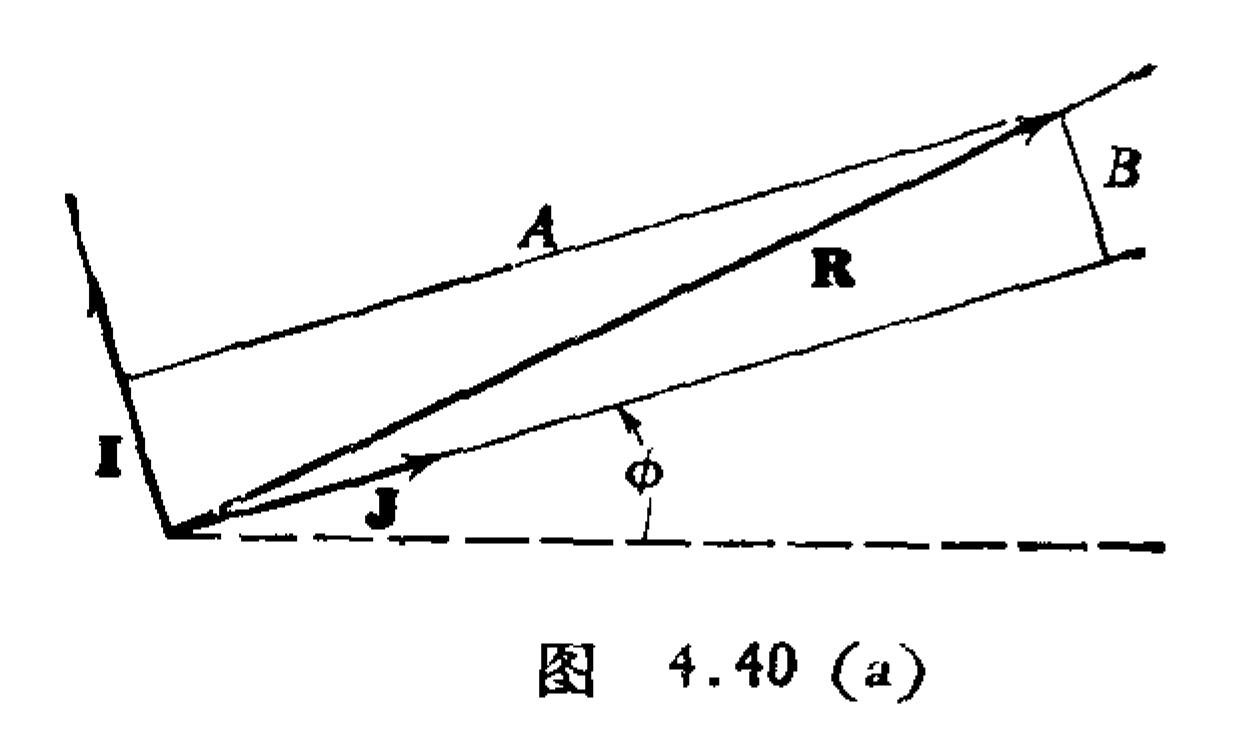


图 4.40 (a)

(a)). 如果这样的 $\mathbf{R}$ 的表示式存在，那么容易求得 A，B 的值. 因为，按假定向量 $\mathbf{I}$ 和 $\mathbf{J}$ 是长为 1 的正交单位向量，所以

$$\mathbf{I}\cdot\mathbf{I}=\mathbf{J}\cdot\mathbf{J}=1,\quad \mathbf{I}\cdot\mathbf{J}=0. \tag{32}$$

分别作出方程(31)和 $\mathbf{I}$, $\mathbf{J}$ 的数量积，就看到 A 和 B 的值必定是

$$A=\mathbf{R}\cdot\mathbf{I},\quad B=\mathbf{R}\cdot\mathbf{J}; \tag{33}$$

换句话说，A，B 是表示 $\mathbf{R}$ 的线段在给定方向上的投影的长度.

如果能够证明 $\mathbf{i}$ 和 $\mathbf{j}$ 本身可以用 $\mathbf{I}$ 和 $\mathbf{J}$ 表示，那么我们就可以由表示式 (29) 推出 $\mathbf{R}$ 可以表示为 $\mathbf{I}$ 和 $\mathbf{J}$ 的线性组合 (31). 现在 $\mathbf{I}=(\alpha,\beta)$，$\mathbf{J}=(\gamma,\delta)$ 可以写成

$$\mathbf{I}=\alpha\mathbf{i}+\beta\mathbf{j},\quad \mathbf{J}=\gamma\mathbf{i}+\delta\mathbf{j}. \tag{34}$$

因为有关系式(32)，量 α, β, γ, δ 必须满足所谓的正交关系

$$\alpha^2+\beta^2=\gamma^2+\delta^2=1,\quad \alpha\gamma+\beta\delta=0. \tag{35}$$

如果我们把方程(34)的第一个乘以 δ, 第二个乘以 β, 再相减，我们有

$$(\alpha\delta-\beta\gamma)\mathbf{i}=\delta\mathbf{I}-\beta\mathbf{J}. \tag{36}$$

类似地

$$(\alpha\delta-\beta\gamma)\mathbf{j}=-\gamma\mathbf{I}+\alpha\mathbf{J}. \tag{37}$$

这里，对于互相垂直的单位向量 $\mathbf{I}$ 和 $\mathbf{J}$ 有

$$(\alpha\delta-\beta\gamma)=\mathbf{I}\times\mathbf{J}=\pm 1, \tag{38}$$

其中取上边的还是取下边的符号取决于 $\mathbf{I}$ 到 $\mathbf{J}$ 旋转 90° 的指向是逆时针还是顺时针. 在取定一个符号的情况下，等式(36)和(37)就用 $\mathbf{I}$ 和 $\mathbf{J}$ 表示了 $\mathbf{i}$ 和 $\mathbf{j}$; 把(36)(37)代入(29)就证明了任意向量 $\mathbf{R}$ 的表示公式(31)成立.

公式(31)也可以解释为在坐标轴分别指向 $\mathbf{I}$ 和 $\mathbf{J}$ 的方向的新坐标系之下向量 $\mathbf{R}$ 的表示式. 同时单位向量的分量是该向量方向角的方向余弦. 令 $\mathbf{I}$ 和 $\mathbf{J}$ 分别有方向角 ϕ 和 ψ. 那么

$$\alpha=\cos\phi,\quad \beta=\sin\phi,\quad \gamma=\cos\psi,\quad \delta=\sin\psi.$$

这里，或 $\psi=\phi+\frac{\pi}{2}$ 或 $\psi=\phi-\frac{\pi}{2}$. 在第一种情况(相当于坐标向量 $\mathbf{I}$, $\mathbf{J}$ 的右手系)，我们有 $\gamma=-\beta$, $\delta=\alpha$, $\alpha\delta-\beta\gamma=+1$

使得

$$\mathbf{I} = (\cos\phi, \sin\phi), \quad \mathbf{J} = (-\sin\phi, \cos\phi). \tag{39}$$

于是给出向量**R**对于坐标向量**I**, **J**的分量的公式(33)具有形式

$$A = a\cos\phi + b\sin\phi, \quad B = -a\sin\phi + b\cos\phi. \tag{40}$$

这些公式表示同一个向量**R**在两个右手坐标系中的分量之间的关系. 这两个坐标系中将一个坐标系的轴旋转角ϕ就得到另一个坐标系. 如果我们假定这两个坐标系有相同的原点O并且**R**为任意一点P的位置向量$\overrightarrow{OP}$, 那么我们由公式(40)可得坐标变换公式, 这些公式在第385页公式(18)中已经给出, 分量a, b和A, B分别为P在两个坐标系中的坐标.

c. 变向量及其导数和积分

我们很自然地要考虑这样一个向量$\mathbf{R}=(a, b)$, 它的分量a, b为变量t的函数, 譬如说$a = a(t)$, $b = b(t)$. 那么, 对任意一个t, 我们有一个向量

$$\mathbf{R} = \mathbf{R}(t) = (a(t), b(t)).$$

我们称$\mathbf{R}(t)$为t的向量函数. 例如, 一个随时间而运动的点的位置向量.

如果当$t \to t_0$时, $a(t)$有极限a^*, $b(t)$有极限b^*, 我们就说当$t \to t_0$时, $\mathbf{R}(t)$有极限$\mathbf{R}^* = (a^*, b^*)$. 在这种情况下, $\mathbf{R}(t)$的长趋于$\mathbf{R}^*$的长, 如$\mathbf{R}^* \neq 0$, $\mathbf{R}(t)$的方向趋于$\mathbf{R}^*$的方向(即**R**的方向余弦趋于$\mathbf{R}^*$的方向余弦). 如果

$$\lim_{t\to t_0} \mathbf{R}(t) = \mathbf{R}(t_0),$$

就是说, 如果**R**的分量是t的连续函数, 那么就说向量$\mathbf{R}(t)$连续地依赖于t. 连续向量的长和方向(假定$\mathbf{R}(t_0) \neq 0$)也随着t连续地变化.

为引入向量的导数, 我们对于参数的两个值t和$t+h$作差商

$$\frac{1}{h}[\mathbf{R}(t+h) - \mathbf{R}(t)] = \left[\frac{a(t+h)-a(t)}{h}, \frac{b(t+h)-b(t)}{h}\right].$$

我们定义 $\mathbf{R}$ 的导数为此差商在 $h\to 0$ 时的极限：

$$\dot{\mathbf{R}}=\frac{d\mathbf{R}}{dt}=\lim_{h\to 0}\frac{1}{h}[\mathbf{R}(t+h)-\mathbf{R}(t)]$$
$$=\left[\frac{da}{dt},\frac{db}{dt}\right]=(\dot{a},\dot{b}).$$

向量的导数由对其分量微分而得到.

容易看出向量乘积的导数满足通常的法则

$$(\mathbf{RS})^{\cdot}=\frac{d(\mathbf{R}\cdot\mathbf{S})}{dt}=\frac{d\mathbf{R}}{dt}\cdot\mathbf{S}+\mathbf{R}\cdot\frac{d\mathbf{S}}{dt}=\dot{\mathbf{R}}\mathbf{S}+\mathbf{R}\dot{\mathbf{S}},$$
$$(\mathbf{R}\times\mathbf{S})^{\cdot}=\frac{d(\mathbf{R}\times\mathbf{S})}{dt}=\dot{\mathbf{R}}\times\mathbf{S}+\mathbf{R}\times\dot{\mathbf{S}},$$

其中对于外积，因子必须取原来的次序.

类似地，我们用 $\mathbf{R}(t)$ 的分量的积分定义向量 $\mathbf{R}(t)$ 的积分：

$$\int_\alpha^\beta\mathbf{R}(t)dt=\left(\int_\alpha^\beta a(t)dt,\ \int_\alpha^\beta b(t)dt\right).$$

由微积分基本定理推出

$$\frac{d}{dt}\int_\alpha^t\mathbf{R}(s)ds=\mathbf{R}(t).$$

d. 对平面曲线的应用．方向，速度和加速度

速度向量

在 4.1 节中，我们使用两个函数 $x=\phi(t)$，$y=\psi(t)$ 表示曲线 C．这些函数的定义域中的每个 t 决定了 C 上的一点 $P=(x,y)$；这里，可以把 t 当成时间，而把 P 当成运动着的点，P 的位置在时刻 t 为 $x(t)$ 和 $y(t)$．如果我们令 x 和 y 为 P 的位置向量 $\mathbf{R}=\overrightarrow{OP}$ 的分量，那么 C 就可由位置向量的终点描绘（图 4.41），

$$\mathbf{R}=\mathbf{R}(t)=(x(t),y(t)).$$

对于相应于 t 和 $t+\Delta t$ 的 C 上的两点 P 和 P'，我们有

$$\overrightarrow{PP'}=\overrightarrow{OP'}-\overrightarrow{OP}=\mathbf{R}(t+\Delta t)-\mathbf{R}(t)=\Delta\mathbf{R}.$$

此向量表示以 P，P' 为端点的 C 的有向割线．这里，如果 Δt 是正的，即如果在 t 增加的方向上 C 上点 P' 在 P 的后面，那么向量

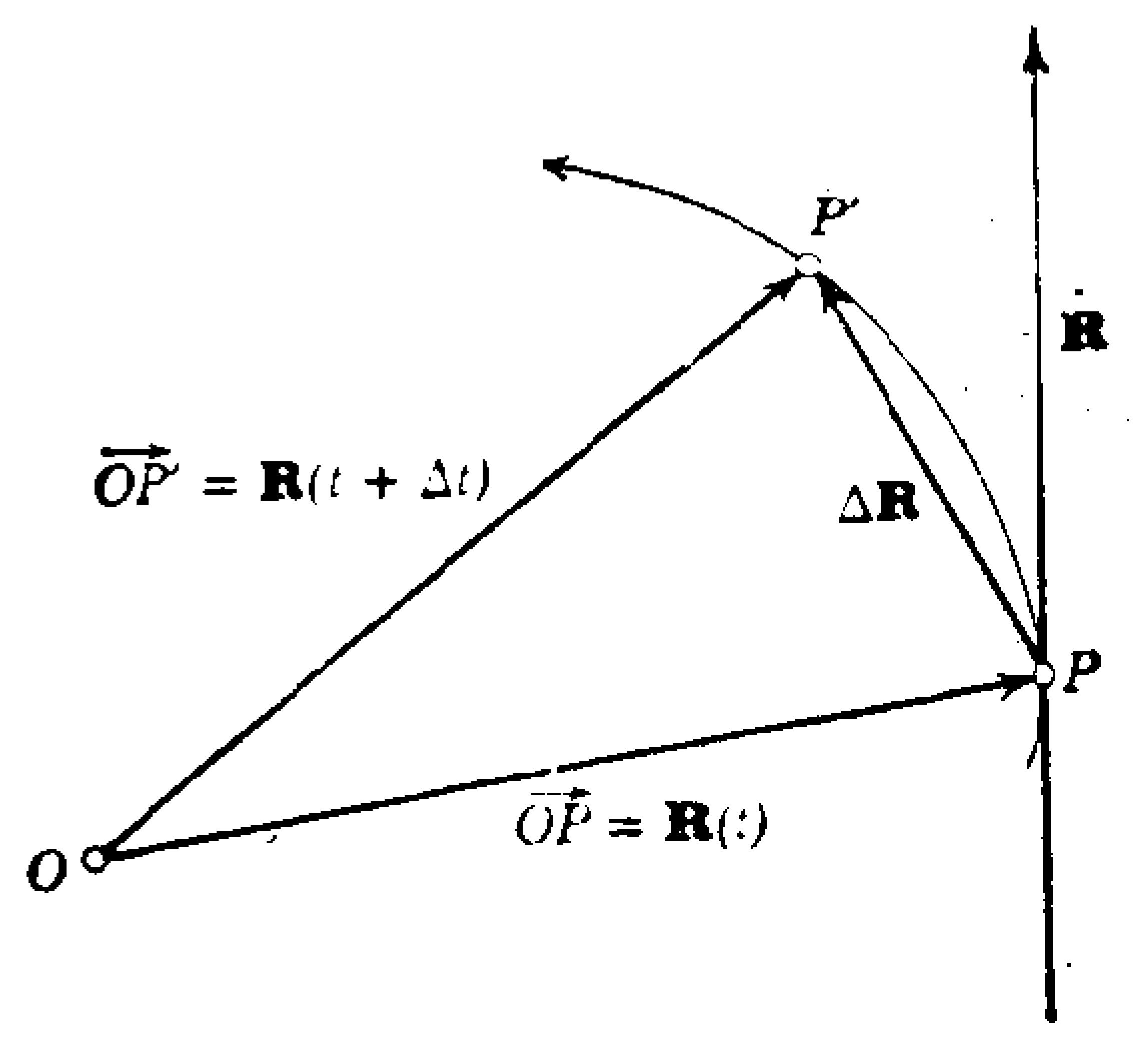


图 4.41 曲线位置向量的导数

$$\frac{\mathbf{R}(t + \Delta t) - \mathbf{R}(t)}{\Delta t}$$

的方向和向量 $\mathbf{R}(t + \Delta t) - \mathbf{R}(t) = \overrightarrow{PP'}$ 的方向一致；它的长是点 P 和 P' 的距离除以 Δt. 当 $\Delta t \to 0$ 时，我们得到极限向量

$$\dot{\mathbf{R}} = \dot{\mathbf{R}}(t) = (\dot{x}(t), \dot{y}(t)),$$

这里再次使用点来表示对参数 t 的导数. $\dot{\mathbf{R}}$ 的方向是割线 PP' 方向的极限，因此是在点 P 的切线的方向. 更确切地，$\dot{\mathbf{R}}$ 指向 C 上相应于 t 增加的切线的方向，条件是 $\dot{\mathbf{R}} \neq 0$. $\dot{\mathbf{R}}$ 的方向余弦与在第 369 页上给出的切线的方向余弦

$$\cos\alpha = \frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}}, \quad \sin\alpha = \frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}}$$

是相同的. $\dot{\mathbf{R}}$ 的长

$$|\dot{\mathbf{R}}| = \sqrt{\dot{x}^2 + \dot{y}^2}$$

可以解释为 $\frac{ds}{dt}$，即沿曲线的弧长 s 对于参数 t 的变化率. 如果 t 表示时间，我们就有 $|\dot{\mathbf{R}}|$ 为点沿着曲线运动的**速率**.

在力学中，质点的速度不仅有确定的**大小**（速率），还要有确定的**方向**. 于是**速度**由**向量** $\dot{\mathbf{R}} = (\dot{x}, \dot{y})$ 表出. $\dot{\mathbf{R}}$ 的长度就是速率，$\dot{\mathbf{R}}$ 的方向就是运动的瞬时方向，即 t 增加的指向的切线方向.

加速度

类似地，质点的加速度定义为向量 $\ddot{\mathbf{R}}=(\ddot{x},\ddot{y})$. 零加速度意味着 $\ddot{x}=\ddot{y}=0$; 如果沿着整个 t 区间 $\ddot{\mathbf{R}}=0$, 那么速度分量为常数值 $\dot{x}=a$, $\dot{y}=b$; 这时，位置向量本身的分量是 t 的线性函数：$x=at+c$, $y=bt+d$. 在这种情况下，质点沿直线以常速率运动.

如果曲线以位置向量 $\mathbf{R}=\mathbf{R}(t)=(x(t),y(t))$ $(\alpha\leqslant t\leqslant\beta)$ 描述，那么我们前面的所有属于曲线的结果很容易用向量符号表示. 对于弧长(见第 373 页式(8))，我们有

$$\int_\alpha^\beta|\dot{\mathbf{R}}|\,dt,$$

而对于曲线包围的有向面积（参见第 389 页，公式(20))，我们有

$$A=\frac{1}{2}\int_\alpha^\beta\mathbf{R}\times\dot{\mathbf{R}}\,dt$$

(这个量的符号仍然依赖于曲线的定向). 最后，对于曲率（参见第 380 页式(15))，我们有

$$\kappa=\frac{\dot{\mathbf{R}}\times\ddot{\mathbf{R}}}{|\dot{\mathbf{R}}|^3}.$$

加速度的切线分量和法线分量

如果我们仍然把 t 理解为时间，那么这些公式有十分有趣的含意. 令 γ 为向量 $\ddot{\mathbf{R}}$ 与向量 $\dot{\mathbf{R}}$ 的夹角. 量 $|\ddot{\mathbf{R}}|\cos\gamma$ 表示 $\ddot{\mathbf{R}}$ 在 $\dot{\mathbf{R}}$ 的方向上的投影；我们称之为加速度的切线分量. 类似地，$|\ddot{\mathbf{R}}|\sin\gamma$ 为 $\ddot{\mathbf{R}}$ 在法线上(更确切地，在由 $\dot{\mathbf{R}}$ 反时针旋转 90° 而得到的法线上)的投影；这就是加速度的法线分量(见图 4.42). 按内积和外积的定义，有

$$|\ddot{\mathbf{R}}|\cos\gamma=\frac{\dot{\mathbf{R}}\cdot\ddot{\mathbf{R}}}{|\dot{\mathbf{R}}|},\quad |\ddot{\mathbf{R}}|\sin\gamma=\frac{\dot{\mathbf{R}}\times\ddot{\mathbf{R}}}{|\dot{\mathbf{R}}|}.$$

现在，

$$\begin{aligned}\dot{\mathbf{R}}\cdot\ddot{\mathbf{R}}&=\frac{1}{2}(\dot{\mathbf{R}}\cdot\ddot{\mathbf{R}}+\ddot{\mathbf{R}}\cdot\dot{\mathbf{R}})=\frac{1}{2}\frac{d}{dt}(\dot{\mathbf{R}}\cdot\dot{\mathbf{R}})\\&=\frac{1}{2}\frac{dv^2}{dt}=v\frac{dv}{dt},\end{aligned}$$

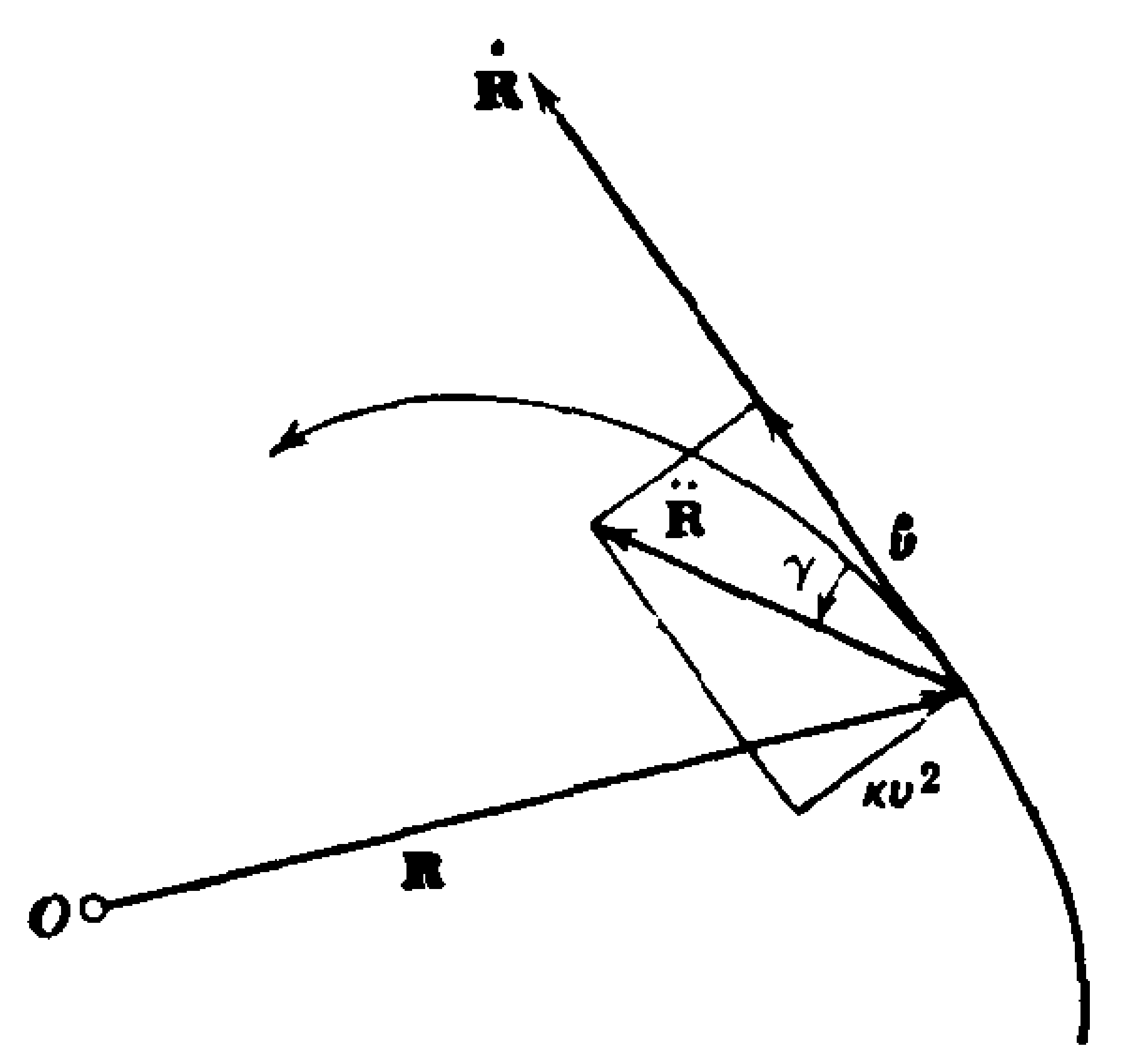


图 4.42　切线加速度和法线加速度

其中 $v=\frac{ds}{dt}=|\dot{\mathbf{R}}|=\sqrt{\dot{\mathbf{R}}\cdot\dot{\mathbf{R}}}$ 是点的**速率**,因此,

$$|\ddot{\mathbf{R}}|\cos\gamma=\frac{dv}{dt}=\dot{v}; \tag{41}$$

所以加速度的切线分量与速率对于时间的变化率是相同的.

对于法线加速度,由曲率公式有

$$|\ddot{\mathbf{R}}|\sin\gamma=\kappa|\dot{\mathbf{R}}|^2=\kappa v^2, \tag{42}$$

即法线加速度等于速率的平方与曲率的乘积.

对于沿着曲线以**常**速率运动的质点,切线加速度 $\dot{v}$ 为零. 因此,加速度向量与曲线垂直. 更确切地,它指向曲线的"内"侧,也就是曲线转向的那一侧(例如这可由以下事实看出,即当 $\kappa>0$ 时,也就是当切线反时针转时, $\sin\gamma>0$). 所以在沿着曲线以常速率运动中,一个点经受着向曲线内侧的加速度,这个加速度与曲率成正比,也与速率的平方成正比. 这个事实有明显的意义,由牛顿定律(后面就要讲到)可知要维持点 P 在此曲线上,就需要有一个与加速度成正比的力.

4.4　在给定力作用下质点的运动

不仅是几何学,而且力学的概念正是在同样程度上决定性地刺激了微积分学的早期发展. 力学所依据的某些基本原理最

初是牛顿建立起来的；这些原理的叙述涉及到导数的概念，而它们的应用需要积分的理论. 这里，我们不详细地分析牛顿的那些原理，而是举一些简单的例子说明微积分在力学中如何应用.

a. 牛顿运动定律

我们只限于考虑单个的质点，即考虑这样的点，质量 m 被想像为集中在这一点. 我们进一步假定运动在 x，y 平面上进行. 在这平面上，质点在时刻 t 的位置由它的坐标 $x=x(t)$，$y=y(t)$ 确定，或等价地由它的位置向量 $\mathbf{R}=\mathbf{R}(t)=(x(t), y(t))$ 确定. 一个量的上方加一点表示对时间 t 的导数. 那么，质点的速度和加速度表示为向量

$$\dot{\mathbf{R}}=(\dot{x}, \dot{y}), \quad \ddot{\mathbf{R}}=(\ddot{x}, \ddot{y}).$$

在力学中，我们把点的运动和作用在该点上的有确定方向和大小的力的概念联系了起来. 力同样以向量 $\mathbf{F}(\rho, \sigma)$ 描述. 作用在同一个质点上的几个力 F_1，F_2，$\cdots$ 的效果和一个单个的力 F，即合力的效果一样. 合力不过是个别的力的向量和 $\mathbf{F}=\mathbf{F}_1+\mathbf{F}_2+\cdots$.

牛顿的基本定律是这样叙述的：质量 m 乘以加速度等于作用在这质点上的力. 用符号表示即

$$m\ddot{\mathbf{R}}=\mathbf{F}. \tag{43}$$

如果我们使用向量的分量写出表达这个基本定律的向量方程，我们就得到等价的一对方程，

$$m\ddot{x}=\rho, \quad m\ddot{y}=\sigma. \tag{44}$$

因为加速度和力仅差一个正因子 m，所以加速度的方向和力的方向相同. 如果没有任何力的作用，即 $\mathbf{F}=0$，那么加速度为零，速度为常数，因而 x 和 y 变为 t 的线性函数. 这就是牛顿的第一定律：如果质点没有受到任何力的作用，它就沿着直线以常速度运动.

牛顿的定律 $m\ddot{\mathbf{R}}=\mathbf{F}$ 起初不过是力的概念的定量定义. 这个关系式的左边可以由对运动的观察结果来确定，然后由此关系得到力.

然而，牛顿定律有更深刻的意义，这是由于在很多情况下，在对相应的运动没有任何了解时，我们可以根据其他的物理考虑确定作用力．这时，这个基本定律就不再是力的定义，而是一种关系，一种我们能够期望由它确定运动的关系．使用牛顿定律的这个决定性的变化在大量例子中，颇起作用，在这些例子中，物理考虑使我们可以把力 **F** 或它的分量 ρ, σ 以明显的方式表示为质点的位置、速度和时间 t 的函数．这时，这个运动定律不是同义语的反复，而给出了用 x, y, $\dot{x}$, $\dot{y}$ 和 t 表示 $m\ddot{x}$, $m\ddot{y}$ 的两个方程，即所谓的运动方程．这些方程是微分方程，即未知函数及其导数的关系式．求解这些微分方程也就是求出所有满足运动方程的各对函数 $x(t)$, $y(t)$，而得到质点在指定的力作用下的全部可能的运动．

b. 落体运动

已知力的最简单的例子是作用在接近地球表面的质点上的重力．由直接观察知道(不考虑空气阻力的影响)每一个落体都有一个加速度，这个加速度垂直向下，并且对所有的物体都有同样的大小 g．用每秒每秒英尺做为测量单位，g 的近似值为 32.16.[1] 如果我们选择 x, y 坐标系，使 y 轴垂直向上，而 x 轴是水平的，那么加速度 $\ddot{\mathbf{R}}=(\ddot{x}, \ddot{y})$ 的分量为

$$\ddot{x}=0, \quad \ddot{y}=-g.$$

这时，根据牛顿基本定律，表示作用在质量为 m 的质点上的重力的向量 **F** 必定是

$$\mathbf{F}=(0, -mg).$$

这个力向量的方向同样是垂直向下的；它的大小，是接近地球表面物体的重量 mg．

当我们消掉因子 m 时，在重力作用下质点运动方程为

$$\ddot{x}=0, \quad \ddot{y}=-g.$$

从这些方程，我们可以容易地得到落体可能有的最一般运动

1) g 的精确值(除了万有引力外，这个值中还包括地球旋转的影响)与在地球上的位置有关．

的描述．对于 t 积分就有

$$\dot{x}=a, \quad \dot{y}=-gt+b,$$

其中 a 和 b 为常数．再一次积分可得

$$x=at+c, \quad y=-\frac{1}{2}gt^2+bt+d,$$

其中 c 和 d 是常数．因此落体的运动方程的一般解依赖四个不定常数 a, b, c, d．我们可以直接地把单个运动的这些常数值与该运动的初始条件联系起来．如果质点在初始时刻 $t=0$ 时位于点 (x_0, y_0)，那么令 $t=0$，我们就有

$$c=x_0, \quad d=y_0.$$

速度 $\dot{\mathbf{R}}=(\dot{x}, \dot{y})=(a, -gt+b)$ 对于 $t=0$ 就成为 (a, b)．因此，(c, d) 和 (a, b) 分别表示质点的初始位置和初始速度．任意选定一组初始条件都唯一地得到一个运动．

在 $a\neq 0$ 时，即在初始速度不是垂直的情况下，我们可以消去 t，而得到质点轨道的非参数表示式．由第一个方程解出 t 来，然后代入第二个方程，就有

$$y=-\frac{g}{2a^2}(x-c)^2+\frac{b}{a}(x-c)+d.$$

因此该路线是抛物线．当 $a=0$ 时，我们有 $x=c=$ 常数，整个的运动沿着一条铅直的直线进行．

c. 约束在给定曲线上的质点的运动

在大多数力学问题中，作用在质点上的力依赖于质点的位置和速度．通常，运动方程太复杂，以致使我们不能确定全部可能有的运动．如果我们可以认为质点描绘的曲线 C 是已知的，那么问题就大大地简化了，只须确定质点沿该曲线的运动．在一大类力学问题中，质点是通过某种机械装置约束在一给定的曲线 C 上运动的．平面摆是一个最简单的例子．把质量 m 用长为 L 的不可伸缩的线与点 P_0 连起来就是平面摆，它在重力影响下，在半径为 L 中心为 P_0 的圆周上运动．

沿曲线 C，我们使用弧长 s 作为参数. 这 时，曲线由 $x=x(s)$，$y=y(s)$ 给出. 那么求质点沿 C 的运动相当于求 s，它作为 t 的函数. 下面给出质点沿该曲线的运动方程.

我们把牛顿公式 $m\ddot{\mathbf{R}}=\mathbf{F}$ 的两边与一个向量 $\boldsymbol{\xi}$ 作内积:

$$m\ddot{\mathbf{R}}\cdot\boldsymbol{\xi}=\mathbf{F}\cdot\boldsymbol{\xi}.$$

如果取 $\boldsymbol{\xi}$ 的长为 1，取 $\boldsymbol{\xi}$ 的方向为 C 的指向 s 增加的切线方向，即 $\boldsymbol{\xi}=\dfrac{d\mathbf{R}}{ds}$，那么在方程 $\mathbf{F}\cdot\boldsymbol{\xi}=f$ 中我们有力的切线分量，或作用在运动方向上的力. 按第 420 页等式(41)，加速度的切线分量 $\ddot{\mathbf{R}}\cdot\boldsymbol{\xi}$ 正好是 $\dfrac{dv}{dt}=\dfrac{d^2s}{dt^2}$，即质点沿曲线的加速度. 这时，牛顿定律就变为公式

$$m\ddot{s}=f, \tag{45}$$

即质点的质量乘以质点沿它的路线的加速度等于沿运动的方向作用在质点上的力.

把这个方程应用到约束在沿曲线 C 上运动的质点，我们假定 f 不包含任何约束力.[1]那么，对于力 $\mathbf{F}=(\rho,\sigma)$，由第 421 页(44)式，我们有

$$f=\rho\frac{dx}{ds}+\sigma\frac{dy}{ds}, \tag{46}$$

因为向量 $\boldsymbol{\xi}$ 有分量 $\dfrac{dx}{ds}$，$\dfrac{dy}{ds}$ (见第 418 页)，对已知的曲线 C，切线的方向余弦 $\dfrac{dx}{ds}$ 和 $\dfrac{dy}{ds}$ 可认为是 s 的已知函数. 如果力 $\mathbf{F}=(\rho,\sigma)$ 同样只依赖于质点的位置，则 f 也是 s 的已知函数. 那么，质点沿 C 的运动就由比较简单的微分方程 $m\ddot{s}=f(s)$ 决定了.

特别，对于重力 $\mathbf{F}=(0,-mg)$，我们有

$$f=-mg\frac{dy}{ds}; \tag{46a}$$

因此约束在曲线 C 上运动的质点在重力影响下的运动方程变为

---

1) 实际上约束的机构必须给出一个力，这个力把质点维持在 C 上 (在单摆的情况下，这个力就由绳子的张力给出). 我们假定这个"反作用"力与曲线垂直，因此没有切线分量；当质点无摩擦地沿一条曲线滑动时就是这样.

$$\frac{d^2s}{dt^2} = -g\frac{dy}{ds}. \tag{47}$$

如果 α 表示曲线的倾角，我们就有 $\frac{dy}{ds} = \sin\alpha$（见图 4.43），运动方程变为

$$\frac{d^2s}{dt^2} = -g\sin\alpha.$$

对于约束在中心在原点，L 为半径的圆周上绕原点运动的质点（单摆），我们有

$$x = L\sin\theta, \quad y = -L\cos\theta,$$

其中 $\theta = \frac{s}{L}$ 为从向下的方向开始计算的极角. 这里（见图 4.44）$\alpha = \theta$，因此

$$\frac{d^2s}{dt^2} = -g\frac{dy}{d\theta}\frac{d\theta}{ds} = -g\sin\theta,$$

或

$$\frac{d^2\theta}{dt^2} = -\frac{g}{L}\sin\theta.$$

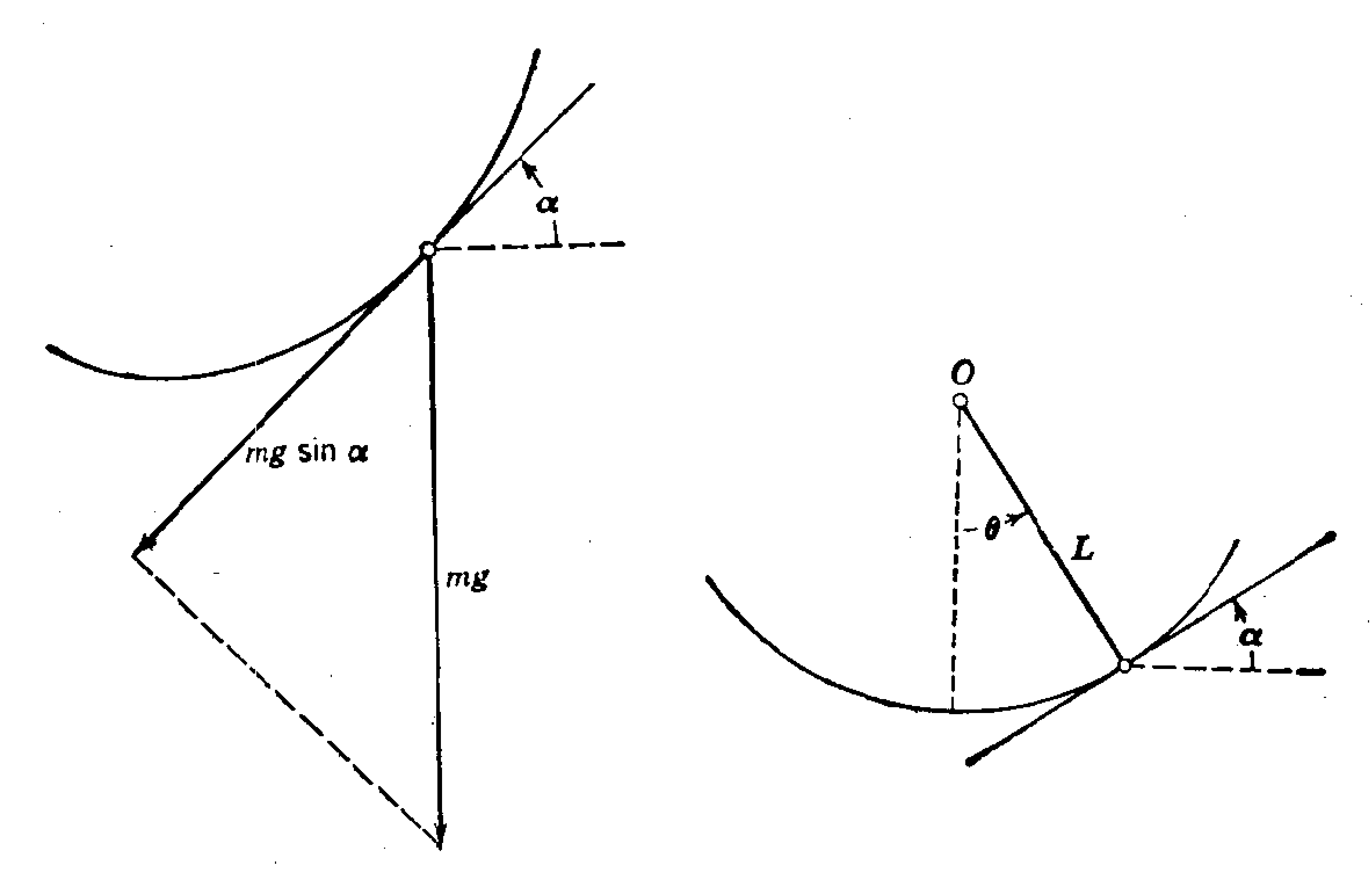


图 4.43 在重力下给定曲线上的运动　　图 4.44 单摆

4.5 受到空气阻力的自由落体运动

我们从质点沿着直线运动的两个例子开始．我们只考虑作用在这条直线的方向上的外力，因而不需要任何约束机构．

自由向下的落体的路线可用参数方程 $x=$ 常数，$y=s$ 来描述．如果作用力仅有重力，我们就有运动方程

$$m\ddot{s} = -mg.$$

如果质点在时刻 $t=0$，从高度 $y_0=s_0$，以初速度 v_0 落下（向上为正值），那么用积分法我们得到

$$s = -\frac{1}{2}gt^2 + v_0t + s_0.$$

如果我们把作用在质点上的摩擦力，即空气阻力的影响也考虑在内，那么我们必须把这个力看成是和运动方向相反的力，并且对于这个力，我们必须作出一定的物理假设[1]．我们将给出不同的物理假设：(a) 阻力和速度成正比，由 $-r\dot{s}$ 的形式给定，其中 r 为正常数；(b) 阻力和速度的平方成正比，对于正的 $\dot{s}$ 有 $-r\dot{s}^2$ 的形式，对于负的 $\dot{s}$ 有 $r\dot{s}^2$ 的形式．按牛顿定律，我们有运动方程

$$m\ddot{s} = -mg - r\dot{s}, \tag{a}$$

$$m\ddot{s} = -mg + r\dot{s}^2, \tag{b}$$

在 (b) 中我们假定了物体正在下落（$\dot{s}<0$）．令 $\dot{s}=v(t)$，我们首先求函数 $v(t)$，那么就有

$$m\dot{v} = -mg - rv, \tag{a}$$

$$m\dot{v} = -mg + rv^2. \tag{b}$$

我们把 t 确定为 v 的函数而不按这些方程把 v 确定为 t 的函数，那么我们的微分方程就可以改写为

$$\frac{dt}{dv} = -\frac{1}{g(1+k^2v)}, \tag{a}$$

1) 这些假设的选择必须适合于所考虑的特殊的物理系统；例如，阻力定律对于低速度和高速度（如子弹速度）是不一样的．

$$\frac{dt}{dv} = -\frac{1}{g(1-k^2v^2)}, \tag{b}$$

其中 $\sqrt{\frac{r}{mg}} = k$. 使用第三章给出的方法，我们可以直接积分而得到

$$t = -\frac{1}{gk^2}\log(1+k^2v)+t_0, \tag{a}$$

$$t = \frac{1}{2gk}\log\frac{1-kv}{1+kv}+t_0. \tag{b}$$

从这些方程解出 v，我们有

$$v = -\frac{1}{k^2}(1-e^{-gk^2(t-t_0)}), \tag{a}$$

$$v = -\frac{1}{k}\frac{1-e^{-2gk(t-t_0)}}{1+e^{-2gk(t-t_0)}}$$

$$= -\frac{1}{k}\tanh[gk(t-t_0)]. \tag{b}$$

这些方程立即显示出运动的一个重要性质. 速度并不是随时间无限增大,而是趋于一个确定的依赖于质量 m 和常数 r（它又依赖于落体的形状和空气密度)的极限. 因为

$$\lim_{t\to\infty} v(t) = -\frac{1}{k^2} = -\frac{mg}{r}, \tag{a}$$

$$\lim_{t\to\infty} v(t) = -\frac{1}{k} = -\sqrt{\frac{mg}{r}}. \tag{b}$$

对于极限速度，摩擦阻力与重力引力刚好平衡. 使用第三章的方法把 $v(t) = \dot{s}$ 的式子再次积分,我们有(可用微分法证实)

$$s(t) = -\frac{1}{k^2}(t-t_0)-\frac{1}{gk^4}(e^{-gk^2(t-t_0)}-1)+c, \tag{a}$$

$$s(t) = -\frac{1}{gk^2}\log[\cosh gk(t-t_0)]+c, \tag{b}$$

其中 c 是积分常数. t_0 是质点在速度为零而高度为 c 的时刻. 如

果把任意时刻 t_1 看成初始条件，我们也很容易把常数 c, t_0 与 t_1 时刻的速度和位置联系起来.

4.6 最简单的一类弹性振动——弹簧的运动

在第二个例子中，我们考虑一质点沿 x 轴被弹性力拉向原点的运动，它有更大的意义. 我们假定弹性力总是指向原点的. 弹性力的大小和质点到原点的距离成正比. 换句话说，我们假设这个力等于 $-kx$，其中系数 k 为弹性连接强度的一个量度. 因为假定 k 为正值，所以当 x 为正时，力是负的；而 x 为负时，力是正的. 牛顿定律告诉我们

$$m\ddot{x} = -kx. \tag{48}$$

这个微分方程本身不能完全地决定这个运动，而对于给定的一瞬间，譬如说 $t=0$，我们可以任意指定初始位置 $x(0)=x_0$ 和初始速度 $\dot{x}(0)=v_0$；用物理的语言就是：我们可使质点从任意的位置，以任意的速度出发；而后运动由微分方程决定. 在数学上，这可由如下事实表达，即微分方程的一般解包含两个不定的积分常数，它们的值我们可用初始条件得到. 我们将马上证明这个事实.

我们能够容易而直接地表示出这个解. 如果我们令 $\omega=\sqrt{\dfrac{k}{m}}$，那么微分方程变为 $\dfrac{d^2x}{dt^2}=-\omega^2x$. 对自变量作替换 $\tau=\omega t$ 就得到第三章第 334 页讨论过的方程 $\dfrac{d^2x}{dt^2}=-x$.因此所有的函数

$$x(t)=c_1\cos\omega t+c_2\sin\omega t$$

满足微分方程，这也可以由微分法证实（其中 c_1 和 c_2 表示任意选择的常数）. 在第三章第 335 页，已经看到我们的微分方程没有任何其他的解，因此每一个在弹性力影响下的这样的运动可由上式给出. 上式也可以容易地表示为

$$x(t)=a\sin\omega(t-\delta)=-a\sin\omega\delta\cos\omega t+a\cos\omega\delta\sin\omega t;$$

只须令 $-a\sin\omega\delta=c_1$, $a\cos\omega\delta=c_2$. 这就引入了新的常数 a 和 δ 代替 c_1 和 c_2. 这种形式的运动叫做**正弦的**或**简谐的**运动. 它

们是周期运动；任何一个状态（即位置 $x(t)$ 和速度 $\dot{x}(t)$）在时间 $T=\frac{2\pi}{\omega}$ 之后重复出现，T 叫做周期，因为函数 $\sin\omega t$ 和 $\cos\omega t$ 有周期 T. 数 a 叫做振动的最大位移或振幅. 数 $\frac{1}{T}=\frac{\omega}{2\pi}$ 叫做振动的频率；它度量了每单位时间振动的次数. 在第八章我们将再来讲振动理论.

*4.7 在给定曲线上的运动

a. 微分方程和它的解

现在我们再回到沿给定的曲线，在任意预先指定的力 $mf(s)$ 作用之下运动问题的一般形式. 我们将根据微分方程（第 424 页方程(45)）

$$\ddot{s}=f(s)$$

来确定 t 的函数 $s(t)$，其中 $f(s)$ 为给定的函数.[1]这个关于 s 的微分方程可用下面的办法完全解出.

我们考虑 $f(s)$ 的任意一个原函数 $F(s)$，因此 $F'(s)=f(s)$. 把方程 $\ddot{s}=f(s)=F'(s)$ 两边乘以 $\dot{s}$. 这样我们可以把左边记为 $\frac{d(\dot{s}^2/2)}{dt}$，因为我们微分 $\dot{s}^2$ 就可以立即看出这一点. 如在 $F(s)$ 中把 s 看作 t 的函数，由锁链微分法则，右边的 $F'(s)\dot{s}$ 为 $F(s)$ 对 t 的导数. 因之，我们应得

$$\frac{d}{dt}\left(\frac{1}{2}\dot{s}^2\right)=\frac{d}{dt}F(s),$$

对上式积分得到

$$\frac{1}{2}\dot{s}^2=F(s)+c,$$

其中 c 表示一个待定常数. 现在，我们得到一个只包含函数 $s(t)$ 和

1) 最初沿曲线的运动方程是 $m\ddot{s}=f(s)$；但是我们总可以把 $f(s)$ 写成 $mf(s)$ 的形式，从而得到这里所用的该方程的较简单的形式.

它的一阶导数的方程(在后面,我们将把这个方程解释为在运动过程中的能量守恒). 我们把方程写成 $\frac{ds}{dt}=\sqrt{2(F(s)+c]}$ 的形式. 我们看到由此并不能立即用积分法直接求出 t 的函数 s. 但是如果我们先满足于求反函数 $t(s)$, 即由质点到达确定的位置 s 所需的时间,那么我们就得到问题的解. 对于 $t(s)$, 我们有方程

$$\frac{dt}{ds}=\frac{1}{\sqrt{2[F(s)+c]}};$$

因此,函数 $t(s)$ 的导数是已知的. 我们有

$$t=\int\frac{ds}{\sqrt{2[F(s)+c]}}+c_1,$$

其中 c_1 为另一个积分常数. 只要我们将最后的积分积出来,我们就解决了问题,因为虽然我们尚未把 s 确定为 t 的函数,但是我们已经反过来求出了时间 t 为位置 s 的函数. 仍然有两个积分常数可利用,这一事实可以使这个一般解适合特定的初始条件.

如果我们令 x 为 s,那么前面的弹性振动例子可以说明这个一般的讨论;这时 $f(s)=-\omega^2s$, 相应地 $F(s)=-\frac{1}{2}\omega^2s^2$. 因此,我们有

$$\frac{dt}{ds}=\frac{1}{\sqrt{2c-\omega^2s^2}},$$

而且

$$t=\int\frac{ds}{\sqrt{2c-\omega^2s^2}}+c_1.$$

引入 $\omega s/\sqrt{2c}$ 作为新变量,能够很容易地求出这个积分: 这样我们有

$$t=\frac{1}{\omega}\arcsin\frac{\omega s}{\sqrt{2c}}+c_1,$$

或给出反函数

$$s=\frac{\sqrt{2c}}{\omega}\sin\omega(t-c_1).$$

这样我们正好得出和前面一样的解公式.

从这个例子，也可以看到积分常数的意义是什么以及如何去确定它们. 例如,如果我们要求在时刻 $t=0$, 质点将在点 $s=0$, 而且在该瞬时有速度 $\dot{s}(0)=1$, 那么我们得到两个方程

$$0=\frac{\sqrt{2c}}{\omega}\sin\omega c_1,\qquad 1=\sqrt{2c}\cos\omega c_1,$$

由此我们得到常数值 $c_1=0$, $c=\frac{1}{2}$. 当任意指定了初始位置 s_0 和初速度 $\dot{s}_0$ (在时刻 $t=0$) 的时候,可以使用完全一样的方法确定积分常数 c 和 c_1.

b. 沿一曲线下滑的质点

使用刚才叙述过的方法，很容易处理在重力作用下质点从没有摩擦的曲线上滑下来的情况. 在第 425 页上我们已经得到了相应于这种情况的运动方程:

$$\ddot{s}=-g\frac{dy}{ds},$$

其中点表示对时间 t 微商. 方程的右边是 s 的已知函数,因为曲线是已知的. 因此,我们把 x 和 y 看作 s 的已知函数.

和上一节一样,我们把方程两边乘以 $\dot{s}$. 左边变为 $\frac{1}{2}\dot{s}^2$ 对 t 的导数. 如果在函数 $y(s)$ 中,我们把 s 看成 t 的函数,那么方程右边为 $-gy$ 对 t 的导数. 因此,求积分,我们有

$$\frac{1}{2}\dot{s}^2=-gy+c,$$

其中 c 为积分常数. 为给出这个常数的解释，我们假定在时刻 $t=0$, 质点是在曲线的坐标为 x_0 和 y_0 的点上,并且假定在这个瞬时质点的速度为 0, 即 $\dot{s}(0)=0$. 那么，令 $t=0$ 我们立即得到 $-gy_0+c=0$, 所以

$$\frac{1}{2}\dot{s}^2=g(y_0-y).$$

因为 $\dot{s}^2$ 不能为负，所以我们看到质点的高度 y 不能超过 y_0 值，只在速度为 0 时，$y = y_0$. 质点越低，速度越大. 现在我们不把 s 看成 t 的函数而考虑反函数 $t(s)$，对此我们立即有

$$\frac{dt}{ds} = \pm \frac{1}{\sqrt{2g(y_0 - y)}},$$

它等价于

$$t = c_1 \pm \int \frac{ds}{\sqrt{2g(y_0 - y)}},$$

其中 c_1 为新的积分常数. 平方根的符号与 $\dot{s}$ 的符号一样，我们注意到如果质点沿着除端点之外处处比 y_0 更低的弧运动，那么符号不能改变. 因为仅当 $\dot{s} = 0$ 亦即 $y - y_0 = 0$ 时才能改变 $\dot{s}$ 的符号. 因此质点只能在曲线最大高度的点上"折回". 代替弧长 s，曲线可以用任意的 θ 作参数，即 $x = \phi(\theta)$，$y = \psi(\theta)$. 引入 θ 作为自变量，我们有

$$t = c_1 \pm \int \frac{ds}{d\theta} \frac{d\theta}{\sqrt{2g(y_0 - y)}} = c_1 \pm \int \sqrt{\frac{x'^2 + y'^2}{2g(y_0 - y)}}\, d\theta,$$

其中，函数 $x' = \phi'(\theta)$，$y' = \psi'(\theta)$ 和 $y = \psi(\theta)$ 为已知的. 为确定积分常数 c_1，我们注意对于 $t = 0$，参数 θ 有值 θ_0. 这就立即给了我们形如下式的解：

$$t = \pm \int_{\theta_0}^{\theta} \sqrt{\frac{x'^2 + y'^2}{2g(y_0 - y)}}\, d\theta. \tag{49}$$

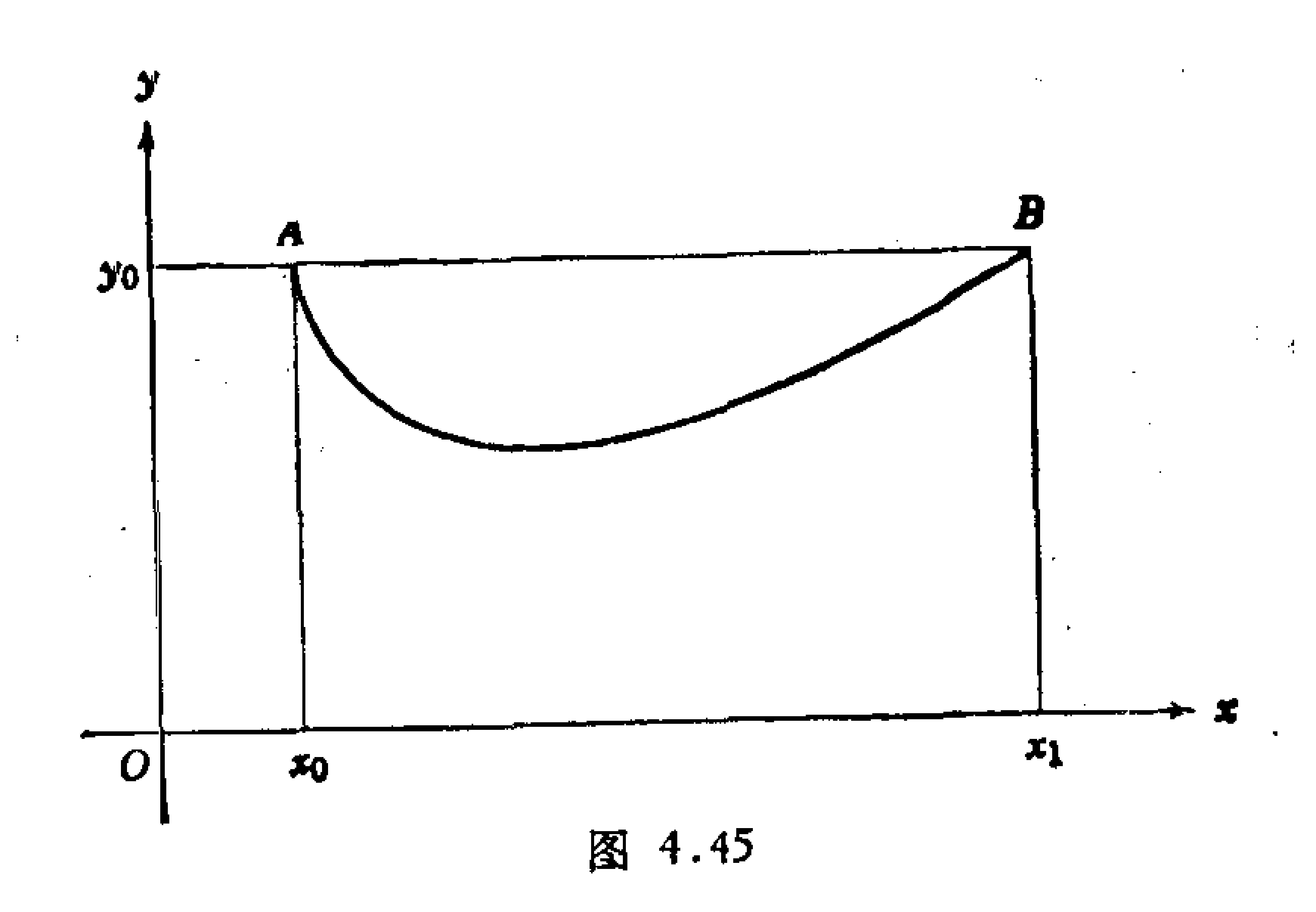


图 4.45

我们看到这个方程表示质点从参数值 θ_0 到参数值 θ 移动所经历的时间。这个函数 $t(\theta)$ 的反函数 $\theta(t)$ 使我们能够完全地描述这个运动；因为在每一个瞬间 t，我们可以确定质点经过的点 $x = \phi[\theta(t)]$，$y = \psi[\theta(t)]$。

c. 运动的讨论

从刚才求得的方程，即使积分的结果没有明显的表达式，我们仍然可以用简单直观的论证推断运动的一般性质。我们假定曲线是图 4.45 所示由向下凸的弧组成的；我们取从左到右为 s 增加的方向。如果质点从对应于 $\theta = \theta_0$，坐标为 $x_0 = \phi(\theta_0)$，$y_0 = \psi(\theta_0)$ 的点 A 开始落下，那么速度是增加的，因为加速度 $\ddot{s}$ 是正值。质点由 A 运动到最低点速度一直是增加的。但是经过最低点之后，加速度变为负的，因为运动方程的右边 $-g\dfrac{dy}{ds}$ 是负的。所以速度是减小的。由方程 $\dot{s}^2 = 2g(y_0 - y)$，我们立即看到当质点到达和初始位置 A 有同样高度的点 B 时，速度的值变为零。因为加速度仍然为负值，所以质点在这个点必定反向运动，使得质点再摆回到原始位置 A。这个动作将无休止地重复进行下去（读者会想到这里已经忽略了阻力）。在这个振动运动中，质点从 B 返回到 A 的时间显然必定和从 A 到 B 的时间相同，因为在相同的高度上，我们有相同的 $|\dot{s}|$ 值。如果我们以 T 表示由 A 到 B，再由 B 返回到 A 的全部路程所需要的时间，那么运动显然是周期为 T 的周期性运动。如果 θ_0 和 θ_1 分别为对应于点 A 和 B 的参数值，那么半周期就有

$$\frac{T}{2} = \frac{1}{\sqrt{2g}}\left|\int_{\theta_0}^{\theta_1}\sqrt{\frac{x'^2 + y'^2}{y_0 - y}}\,d\theta\right|$$

$$= \frac{1}{\sqrt{2g}}\left|\int_{\theta_0}^{\theta_1}\sqrt{\frac{\phi'^2(\theta) + \psi'^2(\theta)}{\psi(\theta_0) - \psi(\theta)}}\,d\theta\right|. \qquad (50)$$

如果 θ_2 是相应于曲线最低点的参数值，那么质点由 A 下落到最低点的时间是

$$\frac{1}{\sqrt{2g}}\left|\int_{\theta_0}^{\theta_2}\sqrt{\frac{x'^2+y'^2}{y_0-y}}\,d\theta\right|.$$

d. 普通摆

最简单的例子是所谓的单摆. 这里所考虑的曲线是有固定半径 L 的圆:

$$x = L\sin\theta,\quad y = -L\cos\theta,$$

其中角 θ 从静止的位置以正指向度量. 由一般的表达式(50), 使用余弦的加法定理我们马上得到

$$T = \sqrt{\frac{2L}{g}}\int_{-\theta_0}^{\theta_0}\frac{d\theta}{\sqrt{\cos\theta-\cos\theta_0}}$$

$$= \sqrt{\frac{L}{g}}\int_{-\theta_0}^{\theta_0}\frac{d\theta}{\sqrt{\sin^2\frac{\theta_0}{2}-\sin^2\frac{\theta}{2}}},$$

其中 $\theta_0(0<\theta_0<\pi)$ 表示单摆振动的振幅, 即质点从时刻 $t=0$, 速度为零开始下落的角的位置.[1]由代换

$$u = \frac{\sin\left(\frac{\theta}{2}\right)}{\sin\left(\frac{\theta_0}{2}\right)},\quad \frac{du}{d\theta} = \frac{\cos\left(\frac{\theta}{2}\right)}{2\sin\left(\frac{\theta_0}{2}\right)},$$

单摆振动的周期的公式变为

$$T = 2\sqrt{\frac{L}{g}}\int_{-1}^{1}\frac{du}{\sqrt{(1-u^2)\left(1-u^2\sin^2\left(\frac{\theta_0}{2}\right)\right)}}.$$

因此我们用椭圆积分(见第 319 页)表示了单摆振动的周期.

如果我们假定振动的振幅很小, 使得我们可以充分精确地把平方根下的第二个因子换成 1, 那么就有

---

1) 这里我们已经假定在运动过程中的某一时刻速度变为零. 这就排除了单摆翻觔斗的运动, 即 θ 不是周期的而是对所有的 t 单调地变化.

$$2\sqrt{\frac{L}{g}}\int_{-1}^{1}\frac{du}{\sqrt{1-u^2}}$$

作为振动周期的近似. 我们可以使用第 280 页积分表中的公式13 求最后这个积分的值，得到 T 的近似值 $2\pi\sqrt{\frac{L}{g}}$. 对于这一级的近似，周期和 θ_0 无关，也就是和摆振动的振幅无关. 显然，精确的周期是比较大的，并且随 θ 增加而增加. 因为在积分区间中

$$1 \geqslant 1 - u^2 \sin^2\frac{\theta_0}{2} \geqslant 1 - \sin^2\frac{\theta_0}{2} = \cos^2\frac{\theta_0}{2},$$

所以我们得到周期的估计值

$$2\pi\sqrt{\frac{L}{g}} \leqslant T \leqslant \frac{1}{\cos\left(\frac{\theta_0}{2}\right)}2\pi\sqrt{\frac{L}{g}}.$$

对于角 $\theta_0 < 10°$，我们有 $\dfrac{1}{\cos\dfrac{\theta_0}{2}} \leqslant \sec 5° < 1.004$，因此周期由 $2\pi\sqrt{\frac{L}{g}}$ 给出，其相对误差小于 0.5%. 关于 T 的椭圆积分的更好的近似可参见 7.6 节 f.

e. 圆滚摆

普通摆振动的周期并不是严格地和摆动的振幅无关的. 这一事实引起了惠更斯 (Huygens) 在制造精确钟的长期努力中寻求一个曲线 C，对于这个曲线，振动的周期与 C 上振动质点开始运动的位置无关.[1]惠更斯发现摆线是这样的曲线.

为了质点实际上能够在摆线上振动，摆线的尖点必须指向重力相反的方向；即我们必须把前面(第 353 页)考虑的摆线绕 x 轴转过来(参见第 353 页图 4.2)，因此我们把摆线方程记为

---

1) 这样的振动称为等时的.

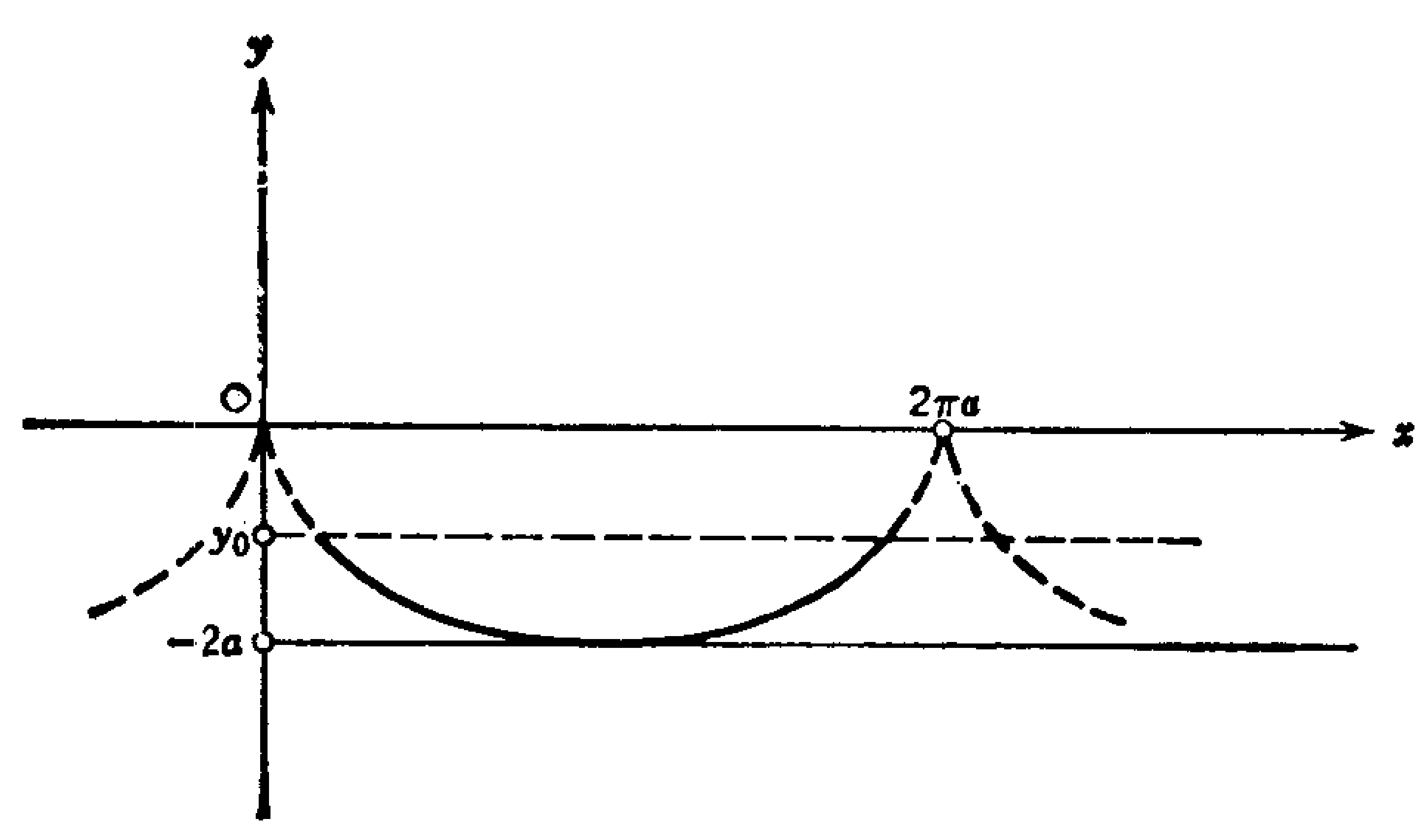


图 4.46 圆滚摆描绘的路线

$$x = a(\theta + \pi + \sin\theta),$$
$$y = -a(1 + \cos\theta),$$

其中包括把参数 t 改变为 $\theta + \pi$（图 4.46）. 质点从高为

$$y_0 = -a(1 + \cos\theta_0) \quad (0 < \theta_0 < \pi)$$

的点向下运动到最低点，再向上到高 y_0 的时间，按第 433 页式(50)是

$$\frac{T}{2} = \sqrt{\frac{1}{2g}} \int_{-\theta_0}^{\theta_0} \sqrt{\frac{x'^2 + y'^2}{y_0 - y}}\, d\theta$$

$$= \sqrt{\frac{2a}{g}} \int_{-\theta_0}^{\theta_0} \frac{\cos\left(\frac{\theta}{2}\right)}{\sqrt{\cos\theta - \cos\theta_0}}\, d\theta.$$

使用和单摆周期刚好相同的替换，我们得到

$$\frac{T}{2} = 2\sqrt{\frac{a}{g}} \int_{-1}^{1} \frac{du}{\sqrt{1 - u^2}},$$

因此我们有

$$T = 4\pi\sqrt{\frac{a}{g}}.$$

所以振动的周期确实与振幅 θ_0 无关. 一个用绳子约束质点使之在摆线上运动的简单方法将在第 451 页叙述.

*4.8 引力场中的运动

作为无约束运动的一个例子，我们考虑在一个吸引物周围的引力场中运动着的一个质点.

a. 牛顿万有引力定律

基于第谷·布拉赫 (Tycho Brahe) 精密的观察，开普勒(Kepler) 描述了行星的运动. 这就导致牛顿发现了任意两个质点之间万有引力的一般定律. 设 $P_0=(x_0, y_0)$ 和 $P=(x, y)$ 是两个质量分别为 m_0 和 m 的质点. 令 $r=\sqrt{(x-x_0)^2+(y-y_0)^2}$ 是质点间的距离. 那么 P_0 作用在 P 上的力 F 有 $\overrightarrow{PP_0}$ 的方向，力的大小 $|F|=\gamma m_0 m/r^2$，其中 γ 为"万有引力常数". 因为 $\mathbf{F}$ 和 $\overrightarrow{PP_0}$ 只差一个正的因子，$\overrightarrow{PP_0}$ 大小为 r，所以我们必定有

$$\mathbf{F}=\frac{\gamma m_0 m}{r^3}\overrightarrow{PP_0}=\left(\frac{\gamma m_0 m(x_0-x)}{r^3},\ \frac{\gamma m_0 m(y_0-y)}{r^3}\right).$$

这个引力定律涉及到质点，也就是涉及到那些物体，它们可以看成集中到一个点上而忽略物体的实际形体(图 4.47). 这样假定的有效性对于天体是完全讲得通的，因为它们相互的距离和它们的直径相比较是惊人地大的. 牛顿大大地扩大了这个定律的应用范围. 他证明了同样的引力定律也可以描述具有相当大体积的质量为 m_0 的物体作用在质量为 m 的质点上的引力，只要物体是一个常密度的球，或者更一般地，物体是由同心的常密度球壳构成的；在这种情况下，物体作用在位于它外边一个质点 P 上的引力很象物体整个质量 m_0 是集中在它的中心 P_0 时对该质点的引力一样(图 4.47). 可以十分准确地认为地球是一常密度的同心球壳. 因此地球作用在它表面上质量为 m 的质点上的引力指向地球的中心 P_0 (即对观察者来说是铅直向下的)，引力的大小为 $\gamma m_0 m/R^2$，其中 R 是地球的半径，m_0 为地球的质量. 这时我们可以令 $\gamma m_0 m/R^2$ 等于 mg，其中 g 为重力加速度(见第 422 页). 换句话说，我们有 $g=\gamma m_0/R^2$.

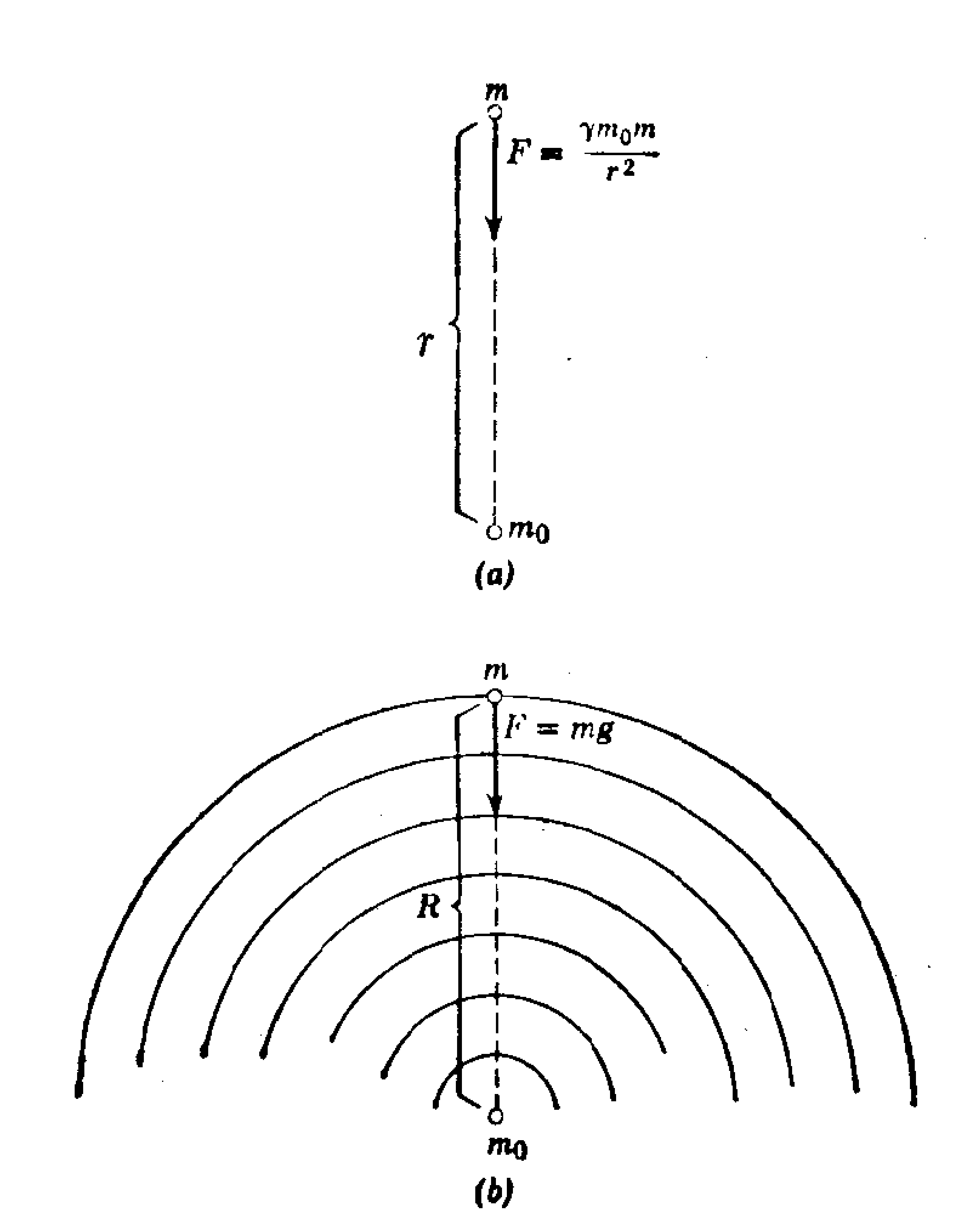


图 4.47 (a) 两个质点的牛顿引力，(b) 地球的重力引力

根据牛顿基本定律，对于质量为 m 的质点 P 在位于 P_0 质量为 m_0 的引力作用下的运动，我们得到运动方程:

$$\ddot{x}=\frac{\gamma m_0(x_0-x)}{r^3},\qquad \ddot{y}=\frac{\gamma m_0(y_0-y)}{r^3}.$$

现在我们进一步做简化假设： m_0 比起 m 要大很多， 以致 P 作用在 P_0 上的引力可以忽略不计，P_0 可以看成是静止的. 例如，这就类似于太阳和行星或地球和它表面上的物体. 我们把坐标原点取在 P_0 点，那么对 $P=(x, y)$ 就有运动方程

$$\ddot{x}=-\frac{\gamma m_0 x}{r^3},\qquad \ddot{y}=-\frac{\gamma m_0 y}{r^3}, \tag{51}$$

其中 $r=\sqrt{x^2+y^2}$.

b. 绕引力中心的圆周运动

我们并不试图得到这些微分方程的最一般的解（众所周知，这

就是相应于沿着圆锥曲线形式的路线的运动．而这个圆锥曲线的一个焦点位于引力的中心)；我们只考虑适合这些方程的最简单的运动类型，就是绕原点的匀速圆周运动和沿着从原点出发的一个半径的运动．对于 P 沿着以原点为中心，a 为半径的圆的匀速圆周运动，我们有 $r=a$ 并且

$$x=a\cos\omega t,\quad y=a\sin\omega t,$$

其中 ω 是常数．运动的周期 T，即在 T 以后 P 又回到相同位置的时间为 $T=2\pi/\omega$．对于速度分量我们有

$$\dot{x}=-a\omega\sin\omega t,\quad \dot{y}=a\omega\cos\omega t,$$

所以 P 在它的轨道上的速率为

$$v=\sqrt{\dot{x}^2+\dot{y}^2}=a\omega=\frac{2\pi a}{T}. \tag{52}$$

P 的加速度分量是

$$\ddot{x}=-a\omega^2\cos\omega t=-\omega^2 x,\quad \ddot{y}=-a\omega^2\sin\omega t=-\omega^2 y.$$

显然，这时它们是满足运动方程(51)的；只要

$$\omega^2=\frac{\nu m_0}{a^3}$$

或

$$a^3=\frac{\gamma m_0}{\omega^2}=\frac{\gamma m_0}{4\pi^2}T^2, \tag{53}$$

这正是**开普勒第三定律**对于圆周运动的特殊情况．按照这个定律，**行星到太阳距离的立方与周期的平方成正比**．

我们可以对开普勒定律给以某些简单的例证．例如，吸引的物体是地球，它的质量为 m_0，半径为 R．注意到这里 $\gamma m_0=gR^2$，我们就有

$$a^3=\frac{gR^2}{4\pi^2}T^2.$$

对于在树顶高度绕地球环行的卫星(当然要忽略空气阻力)，我们有 $a=R\sim 3963$ 英里．根据公式，对于卫星的周期我们有

$$T=2\pi\sqrt{\frac{R}{g}}\sim 1.4\ \text{小时}.$$

对于它在轨道上的速度，则有

$$v = \frac{2\pi R}{T} = \sqrt{Rg} \sim 27,000 \text{ 英尺/秒}. \tag{54}$$

我们可以把绕地球环行的卫星的周期 T 值和月球的周期 27.32 天相比较．月球的周期就是月球回到它相对于星群中原有位置的时间（恒星月）．根据开普勒定律，月球到地球的距离 a 与地球半径之比为它们周期之比的 $\frac{2}{3}$ 次方．我们得到从地球中心到月亮的距离的值为

$$a = \left(\frac{27.32 \times 24}{1.4}\right)^{2/3} R \sim 60R \sim 240,000 \text{ 英里}.$$

这和实际距离的平均值完全一致．

c．径向运动——逃逸速度

我们要考虑的第二种类型的运动是质点由引力中心沿着一条射线，譬如说沿 x 轴的运动．这里，$y=0$，$x=r$，因此运动方程成为

$$\ddot{x} = -\frac{\gamma m_0}{x^2}.$$

按照求解方程 $\ddot{s}=f(s)$ 的一般方法，我们把方程两边乘以 $\dot{x}$，就有

$$\dot{x}\ddot{x} = -\gamma m_0 \frac{\dot{x}}{x^2}$$

或

$$\frac{d}{dt}\left(\frac{1}{2}\dot{x}^2\right) = \frac{d}{dt}\left(\frac{\gamma m_0}{x}\right).$$

由此可见

$$\frac{1}{2}\dot{x}^2 - \frac{\gamma m_0}{x} = h$$

在运动中 h 为常数．（以后我们将知道，这个事实是能量守恒定律的一个特例。）如果把 x 代替 t 作为自变量，那么就有

$$\frac{dt}{dx} = \frac{1}{\dot{x}} = \pm\frac{1}{\sqrt{2h + (2\gamma m_0/x)}},$$

对上式积分可得

$$t = t_0 \pm \int_{x_0}^{x} \frac{d\xi}{\sqrt{2h + (2\gamma m_0/\xi)}}.$$

借助于第三章中叙述的方法，我们可以容易地求出这个积分. 对于在时刻 $t_0 = 0$，在距离 x_0，初速度为零出发的质点我们有 $h = -\gamma m_0/x_0$. 这样的质点落在吸引的质点上（$x = 0$）所需的时间为

$$t = \int_0^{x_0} \frac{d\xi}{\sqrt{2\gamma m_0(1/\xi - 1/x_0)}} = \frac{\pi}{2}\sqrt{\frac{x_0^3}{2\gamma m_0}}.$$

根据开普勒定律，这个时间为距引力中心 x_0 的质点绕引力中心转一圈所需时间的 $\sqrt{\frac{1}{32}}$ 倍（见第 439 页式（53））.

当我们研究质点能够逃逸到无穷远的情况时，关系式

$$\frac{1}{2}\dot{x}^2 - \frac{\gamma m_0}{x} = h$$

有一个有趣的结果. 因为 $\frac{1}{2}\dot{x}^2 \geqslant 0$，所以对于 $x \to \infty$ 我们看到常数 h 必须是非负的，因此在整个运动期间 $\frac{1}{2}\dot{x}^2 - \gamma m_0/x \geqslant 0$. 特别是质点在距离 $x = a$ 处以速度 v 出发要逃逸到无穷远，仅当 $\frac{1}{2}v^2 - \gamma m_0/a \geqslant 0$ 时才有可能. 所以，使质点能逃逸到无穷远的最小可能速度 v 的值为 $v = \sqrt{2\gamma m_0/a}$. 这就是**逃逸速度** v_e. 对于从地球表面出发逃逸到无穷远的质点，即逃脱重力吸引的质点，我们有 $a = R$，$\gamma m_0 = gR^2$，所以

$$v_e = \sqrt{2gR} \sim 37,000 \text{ 英尺/秒}.$$

因此（参见前页公式（54））逃逸速度恰好是把卫星保持在接近地球的圆形轨道上所需速度的 $\sqrt{2}$ 倍. 如果我们忽略空气阻力和地球在它的轨道上的运动，那么流星从无穷远落到地球上撞击时的速度也是 v_e.

4.9 功和能

a. 力在运动中所作的功

功的概念使得前一节的讨论以及力学和物理学中的许多其他问题更加清楚了.

我们仍然设质点在沿着曲线的作用力的影响下在曲线上运动. 我们假定由初始点开始度量的曲线的弧长 s 来确定质点的位置. 这时作用在运动方向上的力本身通常也是 s 的函数. 在力的方向与 s 值增加的方向一致时，这函数的值是正的；在力的方向和 s 值增加的方向相反时，函数的值是负的.

如果沿路径力的大小是常数，那么我们用力与经过的距离 (s_1-s_0) 的乘积来表示力所做的功，s_1 表示运动的终点，s_0 表示起点. 如果力不是常数，我们就使用极限的方法来定义功. 我们把 s_0 到 s_1 的区间分成 n 个相等或不等的子区间. 如果这些子区间很小，那么每一个子区间上的力接近常数；如果 σ_ν 是第 ν 个子区间中任取的一点，那么在这一整个子区间内，力近似地为 $f(\sigma_\nu)$. 如果在整个第 ν 个子区间上的力刚好为 $f(\sigma_\nu)$，那么力所作的功为

$$\sum_{\nu=1}^{n} f(\sigma_\nu)\Delta s_\nu,$$

其中 Δs_ν 照例表示第 ν 个子区间的长. 如果我们取极限，即设 n 无限增加，同时最大子区间的长度趋于零，那么根据积分的定义，我们的和趋于

$$W=\int_{s_0}^{s_1} f(s)ds.$$

我们自然地把它称为力所作的功 (work).

如果力的方向和运动的方向相同，那么力所作的功是正的；我们就说力作了功. 另一方面，力的方向和运动的方向相反，力所作的功是负的；我们就说反抗力作了功[1].

---

1) 这里要注意，我们必须谨慎地描写我们所说的力. 例如，在举重物的过程中，由重力所作的功是负的. 换句话说就是反抗重力作了功. 然而从作上举动作的人的角度看来所作的功是正的，因为人的作用力必须和重力相反.

如果我们把位置的坐标 s 看成时间 t 的函数，因此力 $f(s)=p$ 也是 t 的函数，那么在直角坐标 s 和 p 的平面上，我们就可以把坐标为 $s=s(t)$，$p=p(t)$ 的点画出来．这种点所描绘的曲线，叫做运动的**功图**．如果我们研究在任何机器中的周期运动，那么在一定时间 T（一个周期）之后，运动着的点 $(s(t), p(t))$ 必定回到原来的点；即功图将是闭曲线．在这种情况下，曲线可能仅仅由同一个弧组成，开始向前进行，后来向后进行．例如，发生在弹性振动中的情况就是这样．但是曲线也可能是包围一个面积的更一般的闭曲线；例如，机器的活塞在前进的冲程中受到的压力和后退的冲程中所受的压力是不一样的．那么，在一个循环中，即在时间 T 内所作的功就由功图的负面积给出，换句话说由积分

$$\int_{t_0}^{t_0+T} p(t)\frac{ds}{dt}dt$$

给出，其中从 t_0 到 t_0+T 的时间区间表示运动的一个周期．如果沿正方向经过面积的边界，那么作的功是负的．反之，作的功是正的．如果曲线是由几个回路组成的，经过的方向有些是正的，有些是负的，那么所作的功就由每一个有符号变化的回路面积之和给出．

实际上，这些考虑可用老式的蒸汽机的**示功图**为例来说明．适当地设计一个机械的装置使得铅笔能够在一张纸上移动；铅笔相对于纸的水平运动与从活塞的极端位置到活塞的距离 s 成正比，而铅直地运动与蒸汽压力成正比，因而与蒸汽对于活塞的全部作用力 p 成正比．所以活塞在已知的标度下描绘了蒸汽机的功图．测量这个图形的面积（通常用面积计），就得到蒸汽对于活塞所作的功．这里，我们还看到在第 389 页叙述的面积有符号的规定确实有实际的意义．当蒸汽机轻快地转动时，常常发生这样的情况，在冲程的末端高度膨胀的蒸汽的压力比返回冲程所需要放出蒸汽的压力小；在图上，这个压力以正向通过的回路表示；这时，蒸汽机是从飞轮吸取能量而不是供给能量．

b. 功和动能．能量守恒

运动定律

$$m\ddot{s}=f$$

可以导出质点沿曲线运动时速度的变化和运动方向上的力 f 所作功之间的基本关系．我们使用前面例子中已经用过几次的办法，把运动方程两边乘以 $\dot{s}$：

$$m\dot{s}\ddot{s}=f(s)\dot{s}.$$

现在 $m\dot{s}\ddot{s}=\frac{d}{dt}\left(\frac{1}{2}m\dot{s}^2\right)=\frac{d}{dt}\left(\frac{1}{2}mv^2\right)$，其中 $v(t)=\dot{s}$ 为质点的速度．将方程的两边同时对 t 从 t_0 到 t_1 积分，我们有

$$\frac{1}{2}mv^2(t_1)-\frac{1}{2}mv^2(t_0)=\int_{t_0}^{t_1}f(s)\frac{ds}{dt}dt$$

$$=\int_{s_0}^{s_1}f(s)ds=W.$$

量 $\frac{1}{2}mv^2$ 叫做质点的动能 K．因此，质点在运动中其动能的改变等于作用在质点上的力在运动方向上所作的功．

量 f 表示作用在运动方向上的力或力的切线分量．对于力 $\mathbf{F}=(\rho,\sigma)$，在运动方向上的力是

$$f=\mathbf{F}\cdot\frac{d\mathbf{R}}{ds}=\rho\frac{dx}{ds}+\sigma\frac{dy}{ds}.$$

如果 ρ 和 σ 是 x 和 y 的已知函数，又已知质点沿着曲线 $x=x(s)$，$y=y(s)$ 运动，那么 f 也就成为 s 的已知函数．因此，为计算当质点从一个位置 (x_0,y_0) 移到另一个位置 (x_1,y_1) 时的功

$$W=\int_{s_0}^{s_1}f(s)ds \tag{55}$$

一般地我们必须知道质点运动的路线．

在一类重要情况中，功 W 仅依赖于初始和最终的位置，可以表示为

$$W=V(x_0,y_0)-V(x_1,y_1), \tag{56}$$

决定于一个适当的函数 $V(x, y)$，即势能那么前页上表达动能的改变等于力所作的功的公式也可以记为

$$\frac{1}{2} mv^2(t_1) + V(x_1, y_1) = \frac{1}{2} mv^2(t_0) + V(x_0, y_0). \qquad (57)$$

因此量 $K+V$，动能和势能之和，也就是总的能量在运动过程中不改变．这就是能量守恒这个物理学普遍定律的特例．

在前面讨论过的某些运动中，我们可以容易地构造势能函数 V．例如，对于受到重力作用的质点，我们有 $\mathbf{F}=(0, -mg)$ 和 $f=-mg\frac{dy}{ds}$．所以当质点从位置 (x_0, y_0) 移到位置 (x_1, y_1) 时，重力所作的功是

$$W=\int_{s_0}^{s_1} -mg\frac{dy}{ds}ds=\int_{y_0}^{y_1} -mgdy = mgy_0 - mgy_1.$$

我们看到W和初始与终止位置之间高度的改变成正比．对于势能函数 V，我们可选择 $V=mgy$（或更一般地 $V=mgy+c$，c 为任意常数）．那么，能量守恒定律说，量

$$\frac{1}{2}v^2+gy$$

在运动中是常量．在质点沿曲线滑下的运动的讨论中，我们已经注意到这个事实(第 431 页)．

c. 两个质点间的相互引力

我们可以把势能函数 V 和力联系起来的另一个例子是质量为 m_0 的质点 $P_0=(x_0, y_0)$作用在质量为m的质点 $P=(x, y)$上的重力引力$\mathbf{F}$．这里，

$$\mathbf{F}=\left[\frac{-\mu(x-x_0)}{r^3}, \frac{-\mu(y-y_0)}{r^3}\right],$$

其中 $\mu=\gamma m_0 m$，$r=\sqrt{(x-x_0)^2+(y-y_0)^2}$.（按照库仑定律，同样类型的公式可以给出两个电荷的相互作用.）

由于

$$(x-x_0)\frac{dx}{ds}+(y-y_0)\frac{dy}{ds}=\frac{1}{2}\frac{d}{ds}[(x-x_0)^2+(y-y_0)^2]$$

$$=\frac{1}{2}\frac{dr^2}{ds}=r\frac{dr}{ds},$$

所以，在运动方向上的力是

$$f=-\frac{\mu}{r^3}\left[(x-x_0)\frac{dx}{ds}+(y-y_0)\frac{dy}{ds}\right]$$

$$=-\frac{\mu}{r^2}\frac{dr}{ds}=\frac{d}{ds}\frac{\mu}{r};$$

因此当质点 P 从位置 (x_1, y_1) 移到位置 (x_2, y_2) 时引力所作的功是

$$W=\int_{s_1}^{s_2}\left(\frac{d}{ds}\frac{\mu}{r}\right)ds=\frac{\mu}{r_2}-\frac{\mu}{r_1}=V(x_1, y_1)-V(x_2, y_2),$$

其中 $V(x, y)=-\mu/r=-\mu/\sqrt{(x-x_0)^2+(y-y_0)^2}$ 是势能.

如果我们把质点从位置 (x_1, y_1) 移到无穷远(对应于 $r_2=\infty$)，那么引力所作的功是 $-\mu/r_1$. 使质点移动到无穷远的反向力作的功应是数值相同而符号相反. 因此，$\mu/r_1=-V(x_1, y_1)$ 是为了把质点从位置 (x_1, y_1) 移到无穷远反抗引力所作的功. 这个重要的表达式叫做两个质点的互势. 所以，这里势定义为把两个相互吸引的质点完全分离开所需要作的功. 例如，为了把电子从它的原子中完全分离所需要作的功（电离电势）.

如果把吸引的质量 P_0 视为固定的，那么从能量守恒定律就可得知，被吸引的质点 P 在运动过程中使量

$$\frac{1}{2}v^2-\frac{\gamma m_0}{r}=h$$

(每单位质量 m 的总能量)保持常数值. 我们已经对单纯是径向运动的特殊情况推导过这个事实. 现在我们看到对于在重力引力影响下的任何类型的运动，它也是成立的. 我们可以再次作出结论：要质点逃逸到无穷远必须 $h\geqslant 0$; 它的轨道这时是无界的(抛物线或双曲线)，而不是有界的(椭圆). 相当于 $h=0$ 的逃逸速度

$$v_e = \sqrt{\frac{2\gamma m_0}{r}}$$

是使质点从给定的距离 r 逃逸到无穷远的最小速度．它不依赖于质点出发的方向，而仅依赖于从吸引中心到质点的距离 r．

d. 弹簧的拉伸

第三个例子是拉伸弹簧所作的功．在第 428 页上所作的关于弹簧弹性性质的假定下，作用力是 $f = -kx$，其中 k 是常数．为了把弹簧从未被拉的位置 $x = 0$ 拉伸到最后位置 $x = x_1$ 就必须反抗这个力而作功．因此，这个功由下面的积分给出：

$$\int_0^{x_1} kx\,dx = \frac{1}{2}kx_1^2.$$

***e. 电容器充电**

在物理学的其他分支中可以用类似的方法处理功的概念．例如我们考虑电容器的充电．如果我们用 Q 表示电容器的电量，用 C 表示电容并且用 V 表示电容器的电势差(电压)，那么我们根据物理学知道 $Q = CV$．而且在移动电荷 Q 经过电势差 V 所作的功等于 QV．因为在电容器充电时，电势差不是常数，而是随着 Q 而增加，所以我们可以完全类似于在第 442 页采用的取极限的作法，作为电容器充电所作的功的表达式，我们有

$$\int_0^{Q_1} V\,dQ = \frac{1}{C}\int_0^{Q_1} Q\,dQ = \frac{1}{2}\frac{Q_1^2}{C} = \frac{1}{2}Q_1V_1,$$

其中 Q_1 为已进入到电容器的全部电量，V_1 是在充电过程结束时电容器的电势差．

附　　录

*A.1 法包线的性质

在第 383 页上我们把曲线 C 的法包线 E 定义为 C 的曲率中心

的轨迹. 如果C表示为 $x=x(s)$, $y=y(s)$, 参数 s 是曲线的弧长,那么曲线C的曲率中心(ξ,η)以 s 为参数就是(参见第 383 页(17a))

$$\xi=x-\rho\dot{y},\quad \eta=y+\rho\dot{x},\tag{58}$$

而

$$\kappa=\frac{1}{\rho}=\dot{x}\ddot{y}-\dot{y}\ddot{x}.$$

量κ和$|\rho|$分别为C的曲率和曲率半径.

从这些公式可以推出法包线的一些有趣的几何性质.

对关系式 $\dot{x}^2+\dot{y}^2=1$ 求微商得到 $\dot{x}\ddot{x}+\dot{y}\ddot{y}=0$. 又因为 $\dot{x}\ddot{y}-\dot{y}\ddot{x}=\frac{1}{\rho}$, 所以我们有

$$\ddot{x}=-\frac{1}{\rho}\dot{y},\quad \ddot{y}=\frac{1}{\rho}\dot{x}.\tag{59}$$

将等式(58)对 s 微商得,

$$\dot{\xi}=\dot{x}-\rho\ddot{y}-\dot{\rho}\dot{y}=-\dot{\rho}\dot{y},\quad \dot{\eta}=\dot{y}+\rho\ddot{x}+\dot{\rho}\dot{x}=\dot{\rho}\dot{x},$$

所以

$$\dot{\xi}\dot{x}+\dot{\eta}\dot{y}=0.$$

因为曲线法线的方向余弦是$-\dot{y},\dot{x}$,所以曲线C的法线是法包线E在曲率中心上的切线;或者说法包线的切线是给定曲线的法线;或者说法包线是法线的“包络”(参见图 A.1).

如果我们进一步用σ表示从任意固定点度量的法包线的弧长,用 s 作为参数,那么我们有

$$\dot{\sigma}^2=\left(\frac{d\sigma}{ds}\right)^2=\dot{\xi}^2+\dot{\eta}^2.$$

因为 $\dot{x}^2+\dot{y}^2=1$, 所以由公式(59)得到

$$\dot{\sigma}^2=\dot{\rho}^2.$$

如果我们用适当的方法选择度量σ的方向,那么只要 $\dot{\rho}\neq 0$ 就可得出

$$\dot{\sigma}=\dot{\rho}.$$

积分此式得到

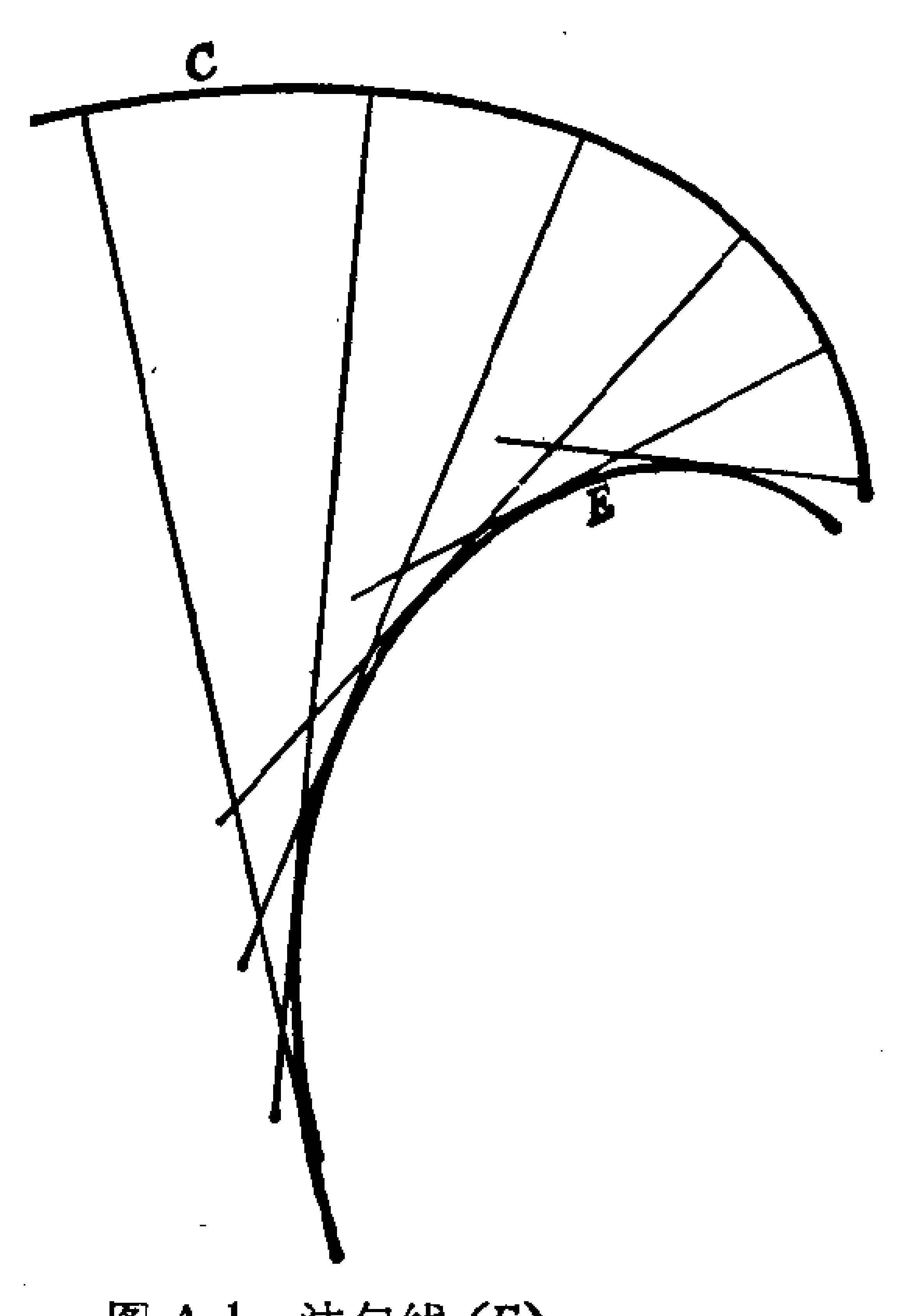


图 A.1 法包线 (E)

$$\sigma_1-\sigma_0=\rho_1-\rho_0.$$

这就是说，法包线两点之间的弧长等于对应的曲率半径之差，条件是对于所考虑的弧 $\dot{\rho}\neq 0$.

这个最后的条件不是多余的. 因为如果 $\dot{\rho}$ 改变符号，那么式 $\dot{\sigma}=\dot{\rho}$ 表明在通过法包线的对应点上弧长 σ 有极大值或极小值；也就是在通过的这个点上我们不能一直继续向前计算 σ，而是我们必须把度量 σ 的指向颠倒过来. 如果我们想避免这种颠倒，我们必须在经过的这一点上把前面公式的符号改变，置 $\dot{\sigma}=-\dot{\rho}$.

也还要注意，对应于曲率半径为极大或极小的曲率中心是法包线的尖点（证明略去）（见图 A.4，A.6）.

用另外的方法也可以表达刚才得到的几何关系：我们想像一条不可伸缩的、可弯曲的线放在沿法包线 E 的弧上，我们把它拉直使得从曲线到法包线的那一部分展开时总是和法包线相切. 如果再加上一个条件，即让这条线的端点开始是在原来的曲线 C 上，那么当我们把这条线转开时，Q 描绘了曲线 C. 这就是渐屈线（展开，转开（evolvere, to unwind））名称的由来. 曲线 C 叫做

渐屈线 E 的渐伸线 (involute). 另一方面，我们也可以从任意曲线 E 着手用这个展开的过程构造出它的渐伸线 C. 那么反过来 E 可看成 C 的渐屈线 (图 A.2).

为了证明这一点，我们考虑曲线 E. 现在它是由 $\xi=\xi(\sigma)$, $\eta=\eta(\sigma)$ 给定的曲线，其中用 ξ 和 η 表示通用直角坐标并且 σ 是 E 的弧长. 如图 A.3 所示，把线拉弯与曲线 E 重合；当线完全地与渐屈线 E 重合时，它的端点 Q 重合于 E 上一点 A 而对应于某一弧长 a. 如果把线转开，直到它在相应的弧长 $\sigma>a$ 的 P 点和渐屈线相切，那么线段 PQ 的长是 $\sigma-a$，它的方向余弦是 $-\dot{\xi}$ 和 $-\dot{\eta}$，其中点表示对 σ 微商. 因此，对于点 Q 的坐标 x, y 我们有

$$x=\xi-(\sigma-a)\dot{\xi},\quad y=\eta-(\sigma-a)\dot{\eta}, \tag{60}$$

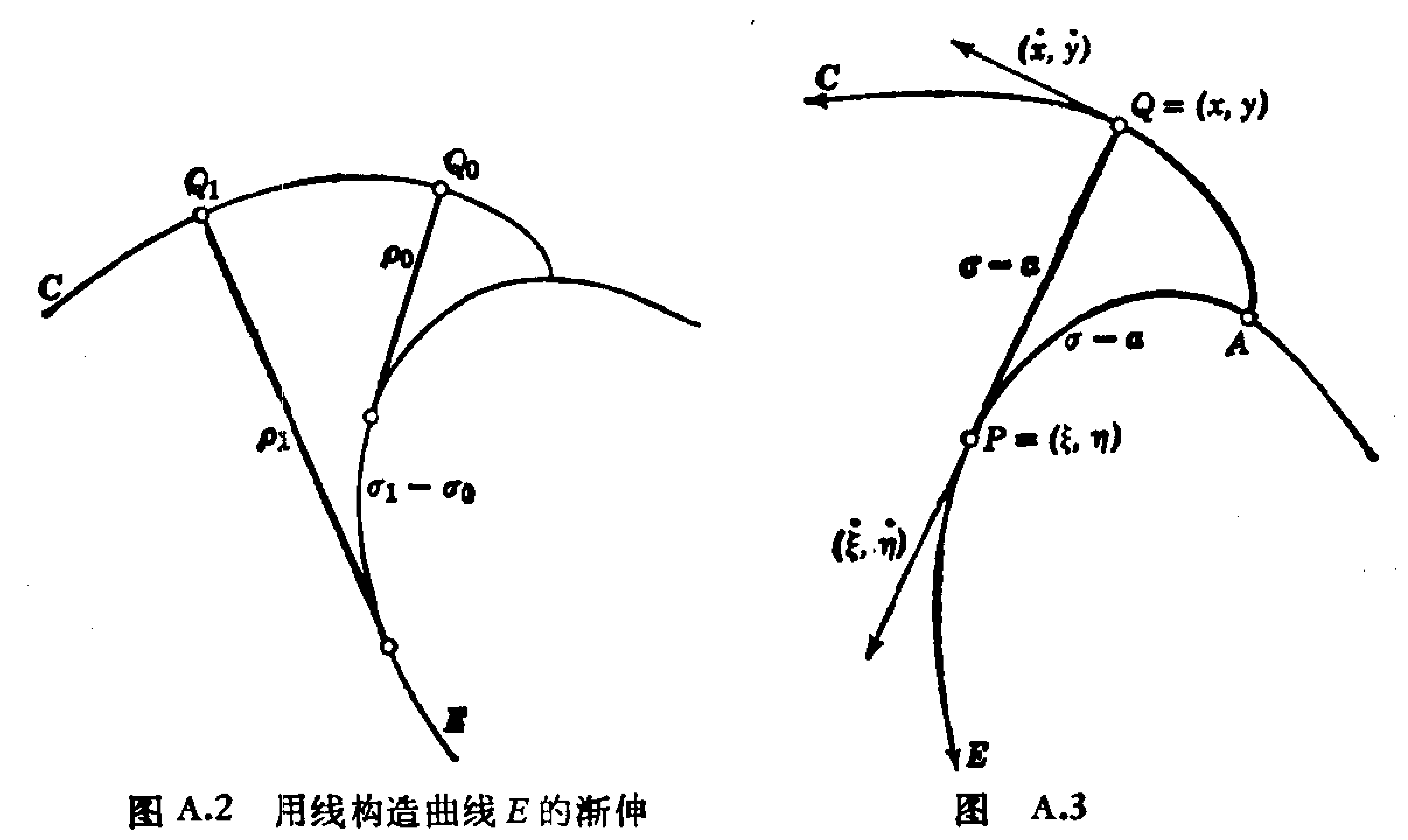


图 A.2 用线构造曲线 E 的渐伸线 C: $\rho_1=\rho_0+\sigma_1-\sigma_0$.

图 A.3

公式(60)给出了用参数 σ 表示的点 Q 描绘的渐伸线的方程. 把它们对 σ 微商我们就得到

$$\begin{aligned}\dot{x}&=\dot{\xi}-\dot{\xi}+(a-\sigma)\ddot{\xi}=(a-\sigma)\ddot{\xi},\\ \dot{y}&=\dot{\eta}-\dot{\eta}+(a-\sigma)\ddot{\eta}=(a-\sigma)\ddot{\eta}.\end{aligned} \tag{61}$$

因为 $\dot{\xi}\ddot{\xi}+\dot{\eta}\ddot{\eta}=0$，所以我们立即有

$$\dot{\xi}\dot{x}+\dot{\eta}\dot{y}=0.$$

这就证明了直线 PQ 是渐伸线 C 的法线，因此我们可以说曲线 C

的法线就是曲线E的切线．因为E的切线的方向余弦是$\dot{\xi}$，$\dot{\eta}$，所以对于C的切线的方向余弦我们有

$$\frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}}=\dot{\eta},\quad \frac{\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}}=-\dot{\xi}. \tag{62}$$

将关系式$\dot{\xi}\dot{x}+\dot{\eta}\dot{y}=0$对$\sigma$微商，然后把由前式(61)(62)得到的$\dot{\xi}$，$\dot{\eta}$，$\ddot{\xi}$，$\ddot{\eta}$代入就有了

$$0=\ddot{\xi}\dot{x}+\ddot{\eta}\dot{y}+\dot{\xi}\ddot{x}+\dot{\eta}\ddot{y}=\frac{\dot{x}^2+\dot{y}^2}{a-\sigma}+\frac{-\ddot{x}\dot{y}+\dot{x}\ddot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}}.$$

因此曲线C对应于点$Q=(x,y)$的曲率半径变为(见第380页公式(15))

$$\rho=\frac{1}{\kappa}=\frac{(\dot{x}^2+\dot{y}^2)^{3/2}}{\dot{x}\ddot{y}-\dot{y}\ddot{x}}=\sigma-a.$$

这也就是从$P=(\xi,\eta)$到点Q的距离．因为P位于C在Q的法线上，所以P为C对应于点Q的曲率中心．因此，**每一个曲线E是它所有的渐伸线的渐屈线**.

例．我们考虑**摆线**

$$x=\pi+t+\sin t,\quad y=-1-\cos t$$

的**渐屈线**．根据式(17)第383页曲线关于任意参数t的曲率中心(ξ,η)是

$$\xi=x-\dot{y}\frac{\dot{x}^2+\dot{y}^2}{\dot{x}\ddot{y}-\dot{y}\ddot{x}},\quad \eta=y+\dot{x}\frac{\dot{x}^2+\dot{y}^2}{\dot{x}\ddot{y}-\dot{y}\ddot{x}}.$$

由简单的计算得到摆线的渐屈线为

$$\xi=\pi+t-\sin t,\quad \eta=1+\cos t.$$

令$t=\tau-\pi$，我们就有

$$\xi+\pi=\pi+\tau+\sin\tau,\quad \eta-2=-1-\cos\tau.$$

这些方程表明，渐屈线本身就是一条类似于原曲线的摆线．如图A.4所示可以把原来的摆线平移而得到渐屈线．

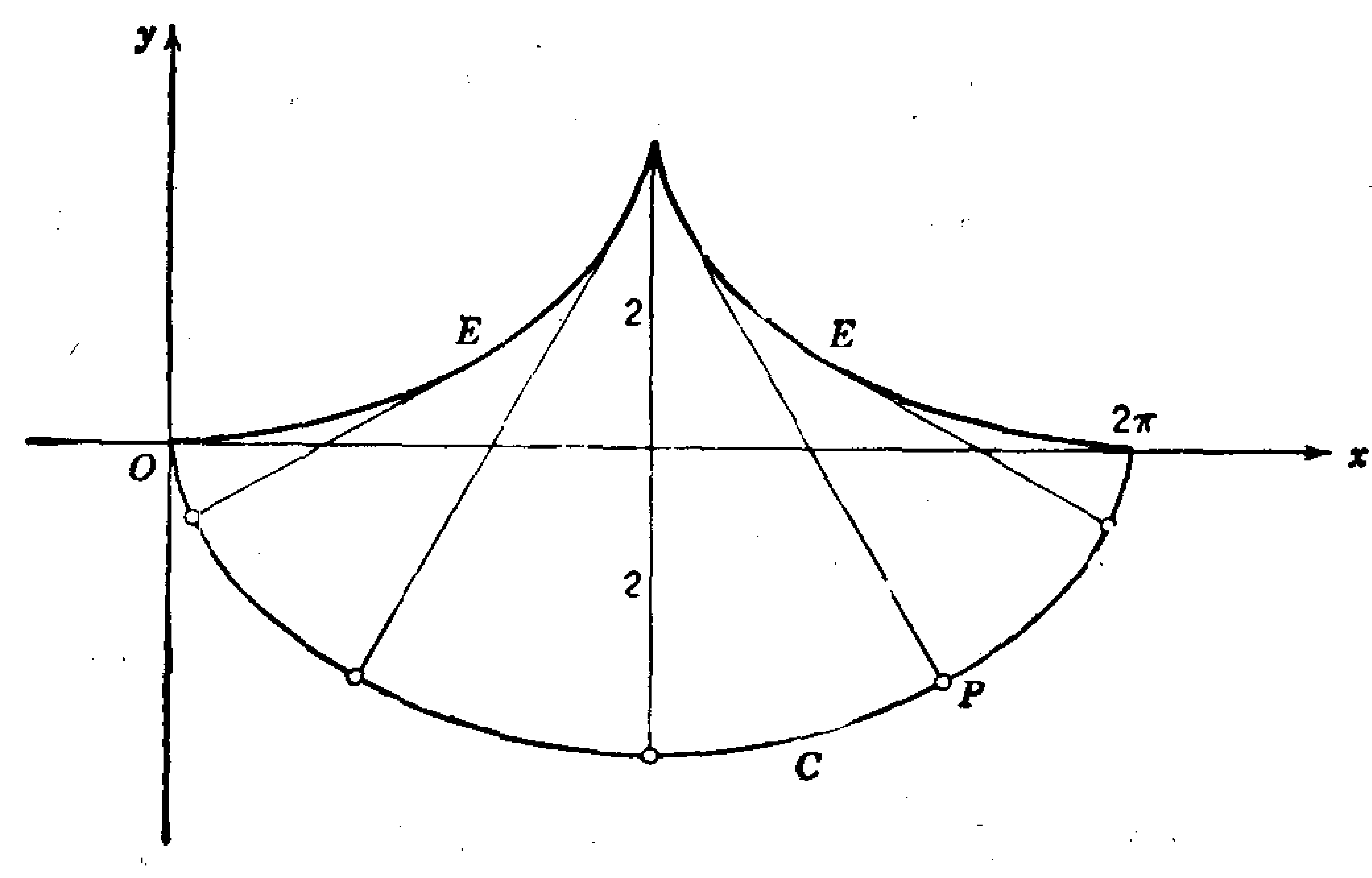
这就给出我们构造**圆滚摆**的一个简单方法(见第435页).如果用一条长为4的线把质量P拴在渐屈线的一个尖点上，那么在张力之下，线的一部分与渐屈线重合，而其余部分位于渐屈线的切线

上．这时质量P位于渐伸线上，即在原来的摆线上．在重力之下，P必定描绘了在摆线某一部分上的等时运动，它的周期与P开始运动的位置无关．（在等时运动中，摆线的参数t并不对应于时间.）在运动过程中，这种类型的摆的自由直线部分的长度是变化的（见图A.4）.

作为进一步的例子，我们推导圆的渐伸线的方程．我们从圆$\xi=\cos\sigma$，$\eta=-\sin\sigma$开始，如图A.5所示，沿切线方向把圆转开．那么给出圆的渐伸线为

$$x=\cos\sigma+\sigma\sin\sigma,\quad y=-\sin\sigma+\sigma\cos\sigma.$$

（使用第450页方程（60），而$a=0$）.



图A.4　圆滚摆

最后我们确定椭圆$x=a\cos t$，$y=b\sin t$的渐屈线．我们立即有

$$\xi=x-\dot{y}\frac{\dot{x}^2+\dot{y}^2}{\dot{x}\ddot{y}-\dot{y}\ddot{x}}=\frac{a^2-b^2}{a}\cos^3 t$$

和

$$\eta=y+\dot{x}\frac{\dot{x}^2+\dot{y}^2}{\dot{x}\ddot{y}-\dot{y}\ddot{x}}=-\frac{a^2-b^2}{b}\sin^3 t$$

作为渐屈线的参数表示式．如用通常的方法从这些方程中消去t，我们就得到非参数形式的渐屈线的方程：

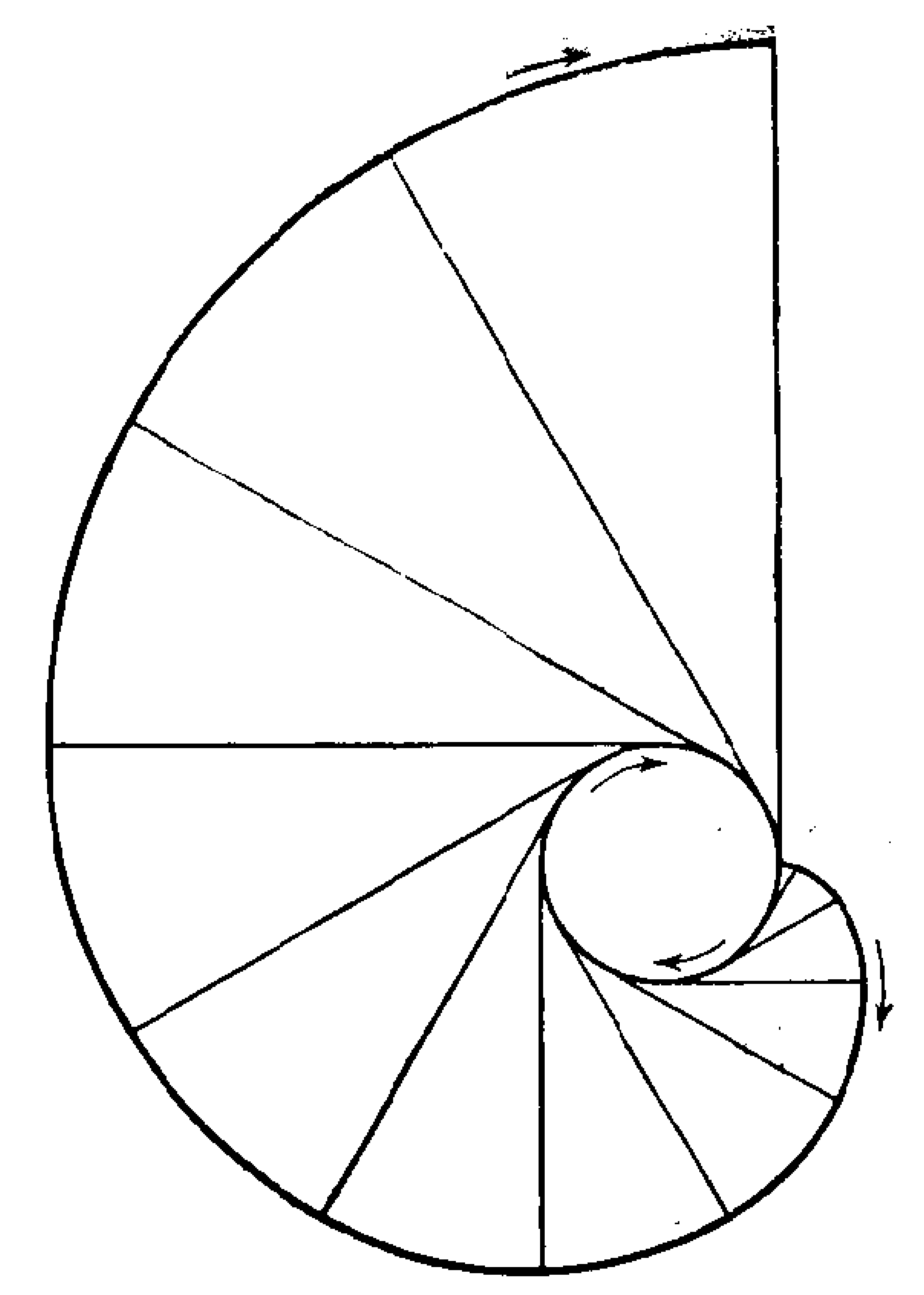


图 A.5　圆的渐伸线.

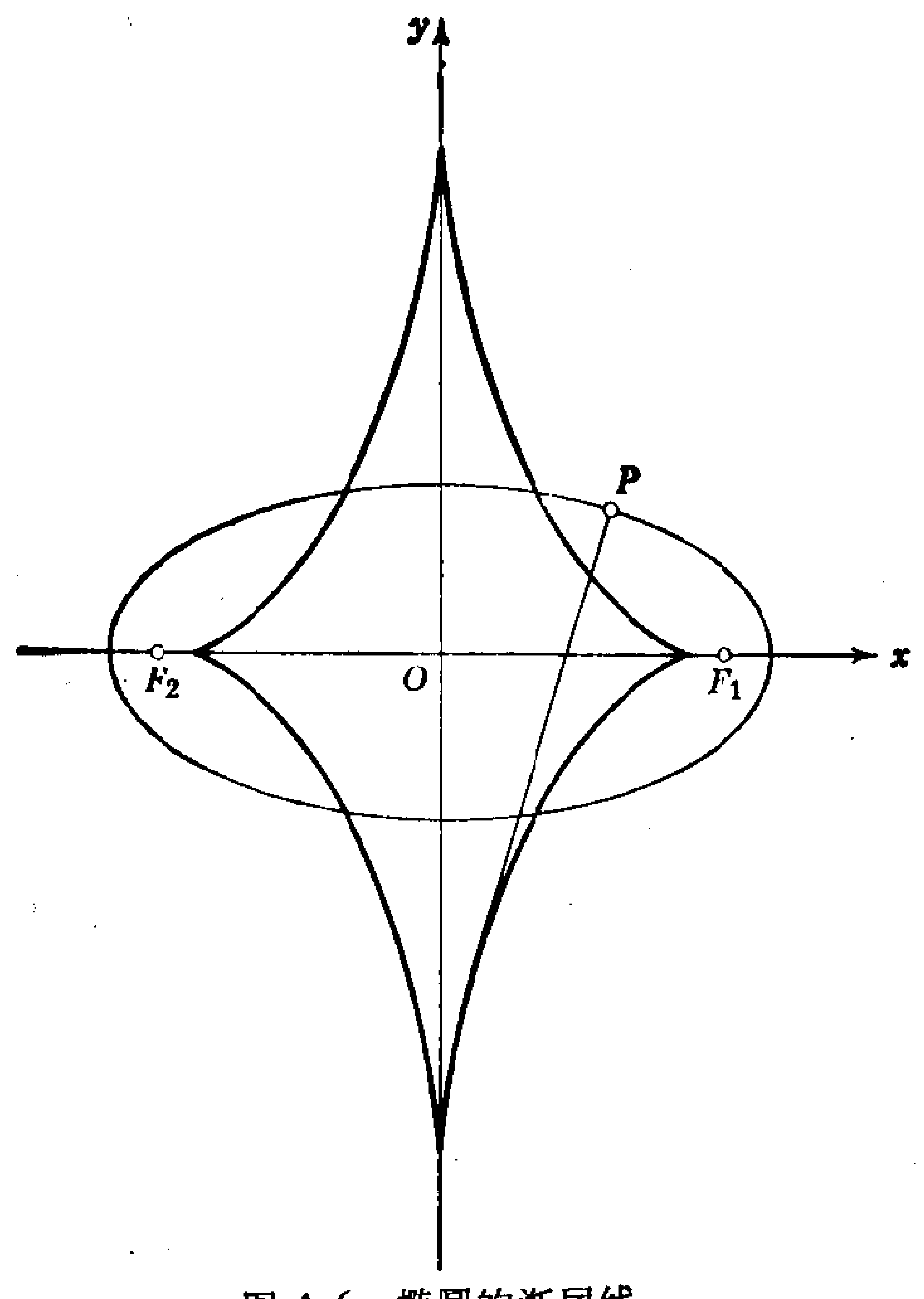


图 A.6　椭圆的渐屈线.

$$(a\xi)^{2/3}+(b\eta)^{2/3}=(a^2-b^2)^{2/3}.$$

这个曲线叫做星形线 (astroid). 图 A.6 给出了它的图形. 由参数方程使我们容易地确信，对应于椭圆顶点的曲率中心实际上是星形线的尖点.

*A.2 闭曲线包围的面积. 指数

在 4.2 节中，处处都不能自己相交的(即所谓的简单的)闭曲线 $x=x(t)$, $y=y(t)$, $\alpha\leqslant t\leqslant\beta$ 包围的有向面积用下面的积分表示，

$$A=-\int_\alpha^\beta y(t)\dot{x}(t)dt;$$

所得到的值是正还是负要依据描绘边界的指向是反时针还是顺时针而定. 如果我们允许曲线可以本身相交，那么这个公式作为 A 的定义仍然是有意义的. 尚待了解的是，在这种情况下 A 如何与面积相联系. 假定曲线 C 的方程为 $x=x(t)$, $y=y(t)$ 并且在有限个点上本身相交，因此把平面分成有限个部分 $R_1, R_2, \cdots$. 进一步假定导数连续，除了或许有限个跳跃不连续点而外；并且 $\dot{x}^2+\dot{y}^2\neq 0$. 最后假定曲线的支线(即 $x=$ 常数的、并且或与曲线相切或通过曲线自身相交的点的竖线)的数目是有限的.

对于每一个 R_i，我们指定由下面方法定义的一个整数，即指数 μ_i: 我们在 R_i 中选择不在任何支线上的任意一点 Q，沿正 y 轴的方向从 Q 点引向上延伸的半直线. 我们计算 t 增加指向的曲线从右到左通过半直线的次数，然后减去曲线 C 从左到右通过半直线的次数；这个差就是指数 μ_i. 例如第 366 页图 4.17 中所示的曲线内部有指数 $\mu=+1$; 在图 A.7 中区域 $R_1, R_2, \cdots, R_6$ 的指数为 $\mu_1=-1$, $\mu_2=-2$, $\mu_3=-1$, $\mu_4=0$, $\mu_5=1$ 和 $\mu_6=0$. 实际上，这个数 μ_i 只依赖于区域 R_i，而不依赖于 R_i 中选择的特殊的点 Q，按下述方式容易看到这一点. 我们在 R_i 中选择另外一个不在支线上的点 Q'，用完全在区域 R_i 中的折线把 Q 和 Q' 连起来 (图 A.8). 当我们沿着这条折线从 Q 前进到 Q' 时，

从右到左通过的次数减去从左到右通过的次数是常数；因为在两根支线之间每一种类型通过的数是不变的，而在越过支线时两种类型的通过数或保持相同，或两种数同时加 1，或两种数同时减 1. 在每一种情况下，差数都是不变的. 这里，我们令支线与曲线相交于几个不同的点，比如说，A，B，$\cdots$，H. 如果 F 竖直地在所有的 A，B，$\cdots$，H 点之下，那么我们可以把 FA，FB，$\cdots$，FH 看成几条不同的支线. 我们前面的论述可应用于每一条这样的线. 因此在确定 μ_i 时不论是用 Q 还是 Q' 数 μ_i 都有相同的值.

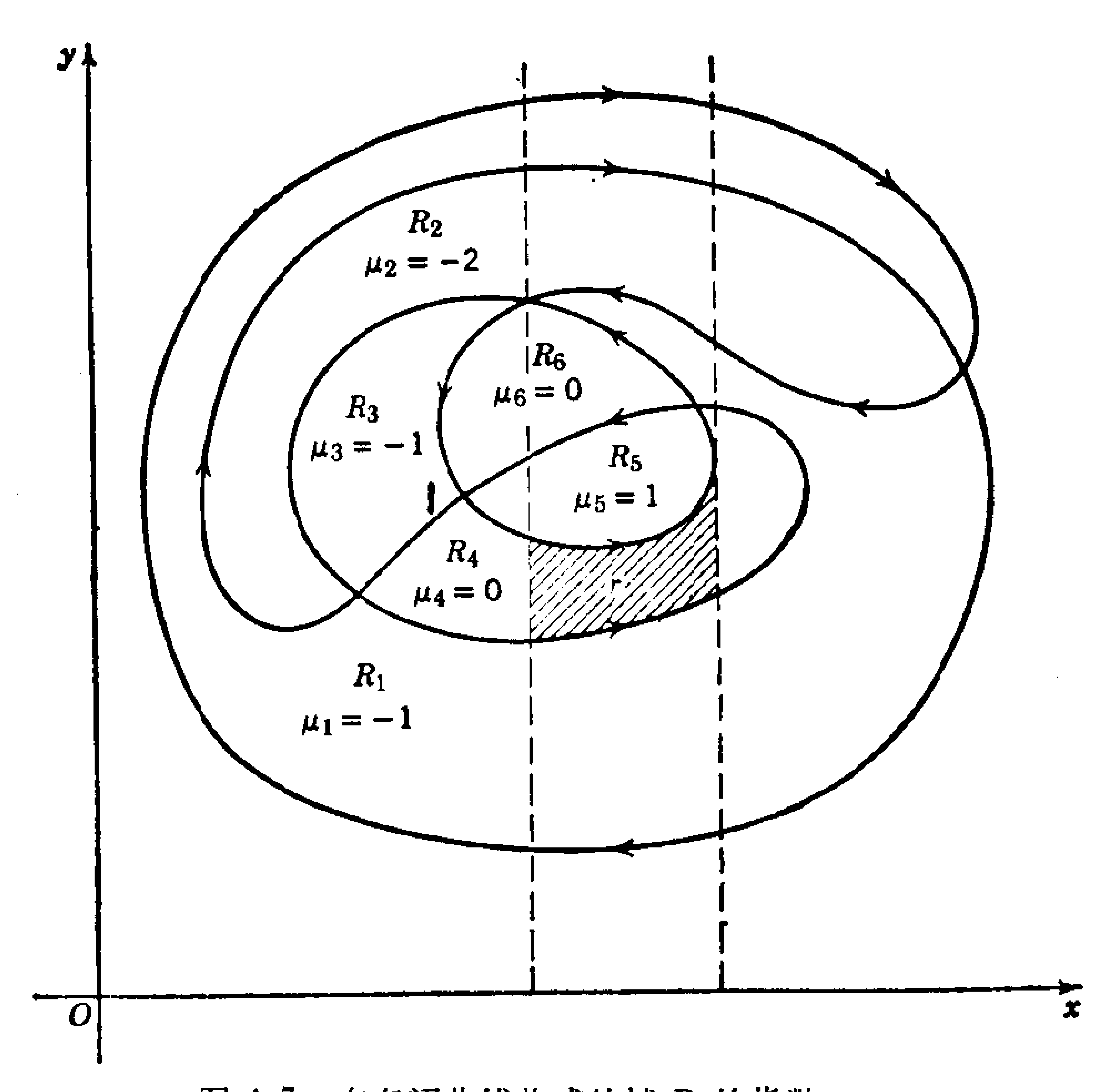


图 A.7　有向闭曲线构成的域 R_i 的指数 μ_i

特别，如果我们的曲线本身不相交，那么曲线的内部由一单个的区域组成，它的指数是 $+1$ 还是 -1 要视描绘边界的指向是反时针还是顺时针而定. 为看清这一点，我们引任意一条竖线(不是支线)与曲线相交. 在这条线上我们找到与曲线相交的最高点 P，在 R 内 P 点以下接近 P 选择一点 Q 使得在 P 和 Q 之间没有任何交点. 那么在 Q 之上曲线有一次通过. 如果曲线以反时针指向通

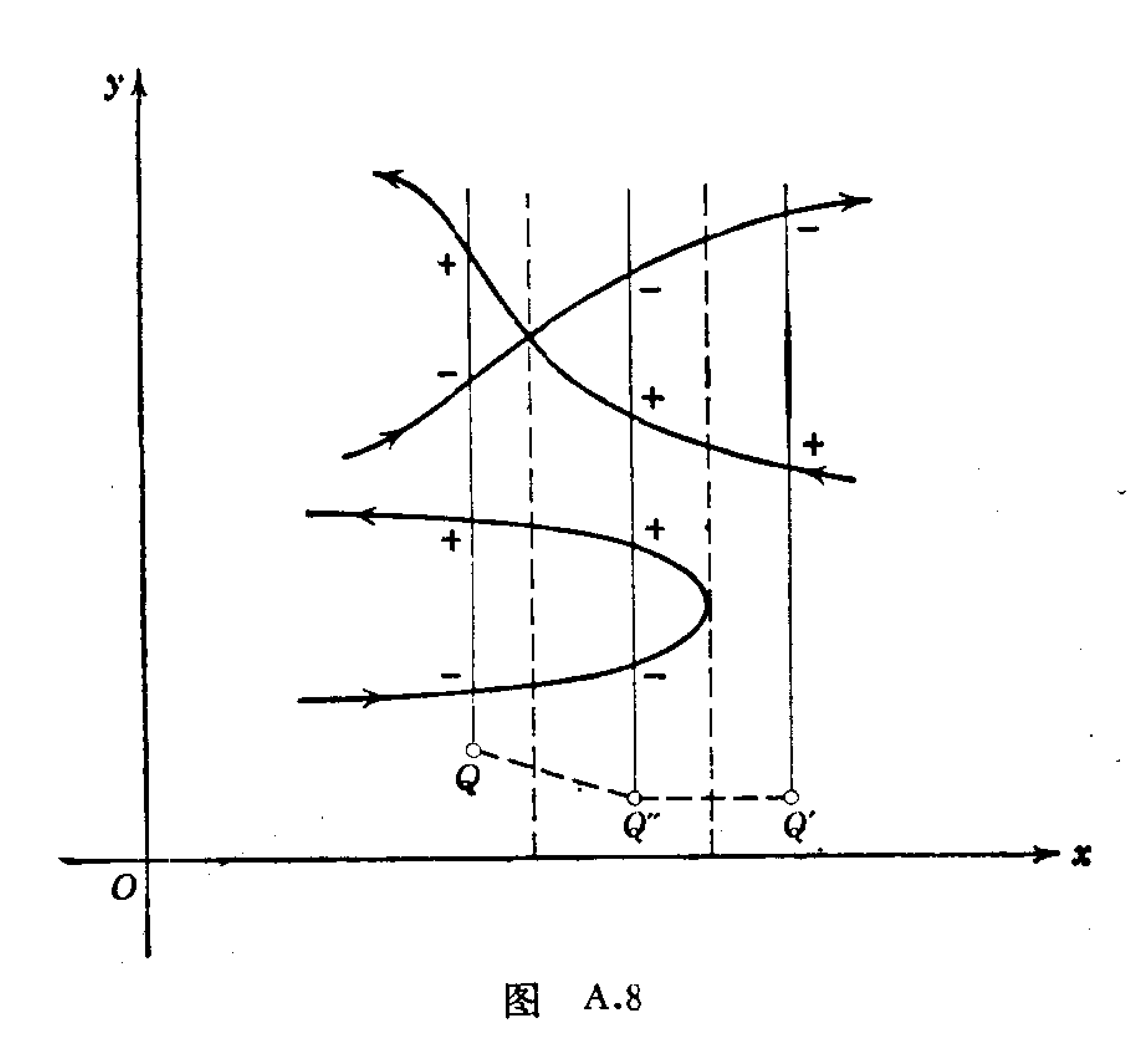


图 A.8

过，那必须是从右到左通过，因之 $\mu=+1$；反之，$\mu=-1$. 正如我们已经看到的，对于R的每一个其他的点保持有相同的μ值. 对于这样的曲线，事实上对于所有的闭曲线有一个区域，即曲线的"外边"，它在所有的方向上无界地伸展；我们立即看到这个区域的指数是零，下面的论述中我们不考虑它. 那么积分A和区域R_i的面积之间的关系由下面的定理给出：

定理. 积分 $-\int_\alpha^\beta y\dot{x}dt$ 的值等于区域 R_i 绝对面积之和，每一个面积 R_i 要计算 μ_i 次；用符号表示即

$$-\int_\alpha^\beta y\dot{x}dt=\sum\mu_i|\text{面积 } R_i|.$$

证明. 证明是简单的. 我们假定（我们有权这样假定）全部曲线位于x轴的上面（对y加一个常数不改变闭曲线的积分A的值）. 支线把R_i分成有限个部分；令r为这些部分之一. 那么，对于函数 $y=y(x)$ 每一个单值的分枝取积分 $-\int y\dot{x}dt=-\int ydx$，并且把它解释为曲线和$x$轴之间的面积. 我们发现，对于在$r$上边的每一个从右到左的分枝要计算 $+1$ 次r的绝对面积，而在r

上边的每一个从左到右的分枝要计算 -1 次；因此，r 的绝对面积总共要计算 μ_i 次．这对于 R_i 的每一个其他的部分同样是正确的．因此，R_i 要计算 μ_i 次．这样，围绕全部曲线的积分值像所说的那样是 $\sum \mu_i|面积R_i|$（参见图 A.7）．这个公式和我们对于简单闭曲线得到的公式是一致的．从我们对曲线 μ 值的讨论中就可看出这一点．

我们对指数 μ_i 的定义有一个缺点，就是它是用特殊的坐标系叙述的．然而，事实上可以证明对于区域 R_i 定义的值 μ_i 只依赖于曲线而和坐标系无关．令曲线 C 上一点沿着 t 增加的方向从 α 到 β 绕着 R_i 中任意固定点 Q_i 以反时针转的总次数（即 C 缠绕 Q_i 的次数）为 ν_i．如我们能验明 μ_i 和 ν_i 相等就证明了 μ_i 与坐标系是无关的．下面我们就证明 μ_i 和 ν_i 相等．

令曲线 C 的参数表示为 $x=x(t)$，$y=y(t)$，其中 $\alpha \leqslant t \leqslant \beta$．设 $Q=(\xi, \eta)$ 为不在 C 的支线上的一点．我们把 Q 取作极坐标系 r，θ 的原点．在这个坐标系中

$$r=\sqrt{(x-\xi)^2+(y-\eta)^2},$$
$$\cos\theta=\frac{x-\xi}{r}, \quad \sin\theta=\frac{y-\eta}{r}.$$

极角确定到只差 2π 的整数倍．然而，如果我们要求 $\theta=\theta(t)$ 沿着曲线 C 对于 t 连续地变化，由 $t=\alpha$ 有 θ_0 值那么 θ 作为 t 的函数就唯一地被确定了．在 $t=\beta$，角 θ 的值是 $\theta(\beta)=\theta_0+2\nu\pi$，其中 ν 是整数．数

$$\nu=\frac{1}{2\pi}[\theta(\beta)-\theta(\alpha)]=\frac{1}{2\pi}\int_\alpha^\beta \frac{d\theta}{dt}dt=\frac{1}{2\pi}\int_C d\theta$$

表示有向曲线 C 缠绕 Q 的次数．

曲线 C 经过从 Q 开始的铅垂半直线是在 $\frac{1}{2\pi}\left[\theta(t)-\frac{\pi}{2}\right]$ 的值为整数 n 的那些 t 值．我们对于固定的 n 考虑参数区间上的 t 值，对于这些 t 值 $\frac{1}{2\pi}\left(\theta-\frac{\pi}{2}\right)=n$．令 σ_n 和 τ_n 为分别有

$\frac{d\theta}{dt}>0$ 和 $\frac{d\theta}{dt}<0$ 这样的 t 值的数．显然，在点 Q 的指数是

$$\mu=\sum_n \sigma_n-\sum_n \tau_n=\sum_n(\sigma_n-\tau_n).$$

另一方面，$\sigma_n-\tau_n$ 只可以取 1，0，−1 中的一个值，因为在 θ，t 平面上 $\theta(t)$ 的图形与直线 $\theta=\frac{\pi}{2}+2n\pi$ 是从上面或从下面交替地相交的．实际上，如果 $\frac{\pi}{2}+2n\pi$ 位于 $\theta(\alpha)$ 和 $\theta(\beta)$ 之间，我们就有 $\sigma_n-\tau_n=\operatorname{sign}[\theta(\beta)-\theta(\alpha)]$，否则 $\sigma_n-\tau_n=0$.

因此，μ 等于在 $\theta(\alpha)$ 和 $\theta(\beta)$ 之间出现 θ 值为 $\frac{\pi}{2}+2n\pi$（n 为整数）的数目按 $\operatorname{sign}[\theta(\beta)-\theta(\alpha)]$ 计算的次数．即 μ 等于数 ν.

因为 $\theta=\arctan[(y-\eta)/(x-\xi)]$，所以我们有

$$\frac{d\theta}{dt}=\frac{\dot{y}(x-\xi)-\dot{x}(y-\eta)}{(x-\xi)^2+(y-\eta)^2}.$$

由此得到有向闭曲线 C 关于点 (ξ,η) 的指数 μ 的积分表示式

$$\mu=\frac{1}{2\pi}\int_\alpha^\beta \frac{\dot{y}(x-\xi)-\dot{x}(y-\eta)}{(x-\xi)^2+(y-\eta)^2}dt,$$

这可以不直接指明参数 t 而简单地写成（见第 391 页）

$$\mu=\frac{1}{2\pi}\int_C \frac{(x-\xi)dy-(y-\eta)dx}{(x-\xi)^2+(y-\eta)^2}.$$

这些结果的值得注意的特点在于，我们可以由 C 的参数表示用求积分的办法来解析地确定描述点 Q 和曲线 C 之间的拓朴关系的整数值 μ 和 ν.

问　　题

4.1c 节第 352 页

1．简略地画出 $a=4c$ 的内摆线（星形线）并求它的非参数方程．

2．证明：如果 $\frac{c}{a}$ 是有理数，那么在动圆滚了整数次以后一般的内摆线

是闭的，而如果 $\frac{c}{a}$ 是无理数，那么曲线与固定圆的圆周有无穷多个交点，而不是闭的。

3. 导出次摆线的参数表示

$$x = at - b\sin t,\quad y = a - b\cos t.$$

次摆线就是缚在半径为 a 的圆盘子上面一点 P 在圆盘沿直线滚动时的路线。圆盘中心到 P 的距离为 b（见图 4.7）。

4. 求曲线 $x^3 + y^3 = 3axy$ 的参数方程(笛卡儿叶形线)，选由原点到点 (x, y) 的射线和 x 轴夹角的正切作为参数 t。

4.1e 节第 367 页

1. 在交点上两曲线之间的夹角 α 定义为在该点它们的切线之间的夹角。使用曲线的参数表示求 $\cos\alpha$ 的公式。

2. 设 $x = f(t)$，$y = g(t)$。导出 $\frac{d^2y}{dx^2}$ 和 $\frac{d^3y}{dx^3}$ 用对 t 的导数表示的公式。

3. 求极坐标中两条曲线 $r = f(\theta)$ 和 $r = g(\theta)$ 之间的夹角的公式

4. 求处处与过原点的直线有相同交角 α 的曲线的方程。

5. 证明：如果 $x = f(t)$ 和 $y = g(t)$ 在闭区间 $[a, b]$ 上连续，在开区间 (a, b) 内可微，且 $x'^2 + y'^2 > 0$，那么在开弧 $x = f(t)$，$y = g(t)$，$(a < t < b)$ 上至少有一点使得该点的切线与连接端点的弦平行。

6. 设 P 为圆上的一点。当圆沿给定的直线滚动时，P 画出摆线。设 Q 是圆与该直线的切点。证明，摆线在 P 点处的法线在任意时刻都经过 Q 点。对于 P 的切线有什么类似的性质?

7. 证明星形线 $x = 4c\cos^3\theta$，$y = 4c\sin^3\theta$ 的切线被坐标轴切断的那部分的长是常数。

*8. 证明两族椭圆和双曲线 $(0 < a < b)$

$$\frac{x^2}{a^2 - \lambda^2} + \frac{y^2}{b^2 - \lambda^2} = 1,\quad 0 < \lambda < a,$$

$$\frac{x^2}{a^2 - \tau^2} + \frac{y^2}{b^2 - \tau^2} = 1,\quad a < \tau < b$$

是共焦的(即有共同的焦点)，并且相交成直角。

9. (a) 证明对于椭圆由两个焦点到曲线的一点所引的两射线之间的角

被该点的法线平分。

(b) 证明对于双曲线上述的交角由切线平分。

4.1f 节 第 371 页

1. 证明由

$$y=\begin{cases} x^2\sin\frac{1}{x}, & 0<x\leqslant 1\\ 0, & x=0\end{cases}$$

定义的曲线有有限长，而由

$$y=\begin{cases} x\sin\frac{1}{x}, & 0<x\leqslant 1\\ 0, & x=0\end{cases}$$

定义的连续曲线是不可求长的。

2. 证明，如果函数 f 在闭区间 $[a, b]$ 上有定义并且单调，那么由 $y=f(x)$，$(a\leqslant x\leqslant b)$ 定义的弧是可求长的。

4.1g 节第 376 页

1. 第二种椭圆积分形为

$$\int_0^\phi\sqrt{1-k^2\sin^2\theta}\,d\theta.$$

(a) 证明可以使用第二种椭圆积分表示椭圆 $x=a\cos\theta$，$y=b\sin\theta$ 的弧长。

(b) 证明次摆线 $x=at-b\sin t$，$y=a-b\cos t$ 的弧长也可用椭圆积分表示。

*(c) 证明可以使用第一种和第二种椭圆积分表示双曲线的弧长。

4.1h 节第 378 页

1. 设 P 为生成摆线的滚动圆上的一点。设 Q 为该圆在任意给定时刻的最低点。证明，Q 平分连结 P 到摆线在 P 点的密切圆中心的线段。

2. 求在 $x=0$ 时，$y=x^2$ 的曲率中心。确定曲线在 $x=0$ 和 $x=\varepsilon$ 时的法线与曲线的交点。计算曲率中心到交点的距离。给出另一个曲率中心的定义，并证明这个定义和本书给出的定义是等价的。

3. 考虑密切圆在切点会不会越过曲线的问题。

*4. 证明曲线 C 在 P 点的曲率圆为过三点 P, P_1, P_2 的圆，当 P_1 趋于 P, P_2 趋于 P 时的极限.

5. 设 $r=f(\theta)$ 是极坐标中曲线的方程. 证明曲率由下式给出:

$$\kappa=\frac{2r'^2-rr''+r^2}{(r'^2+r^2)^{3/2}},$$

其中

$$r'=\frac{df}{d\theta},\quad r''=\frac{d^2f}{d\theta^2}.$$

6. 如果曲线的切线被切点和 y 轴之间截取的长总是等于 1，那么曲线称为曳物线. 试求它的方程. 证明曲线每一点的曲率半径和曲线上该点与 y 轴之间截取的法线长成反比. 计算曳物线的弧长并且求出它的用弧长表示的参数方程.

7. 设 $x=x(t)$, $y=y(t)$ 为一闭曲线. 沿曲线的法线测出一固定的长 p. 那么由这个线段的终点描绘的曲线叫做原来曲线的平行曲线. 试求平行曲线的面积、弧长和曲率半径.

8. 证明曲率为固定的常数 k 的曲线只能是半径为 $\frac{1}{k}$ 的圆.

*9. 如果 xy 平面上曲线的曲率是弧长的单调函数，证明曲线不是闭的并且没有任何二重点.

4.1i 节第 384 页

1. 证明曲线 $x=x(t)$, $y=y(t)$ 的曲率表达式经过坐标轴的旋转是不变的，并且用 $t=\varphi(\tau)$ (其中 $\varphi'(\tau)>0$) 来改变参数时曲率也是不变的.

4.3d 节第 417 页

1. 证明如果加速度总是垂直于速度，则速率是常数.

2. 把速度向量视为位置向量而描绘的曲线称为速端线. 说明沿闭曲线运动的质点是否可能有一条直线作为它的速端曲线.

3. 假定一个滚动圆以常速率运动，试求在这运动中生成摆线的点 P 的速度和加速度.

4. 设 A 是平面上的一个定点，并且假设一个动点 P 的加速度向量总是指向 A 而且正比于 $1/|AP|^2$. 证明速端线是一个圆周 (参考问题 2).

5. 设 A 是一个圆周上的固定点. 设 P 是此圆周上一个动点，其加速度向量指向 A. 证明加速度正比于 $|AP|^{-5}$.

4.5节 第 426 页

1. 沿直线运动的质点受到产生减速度 ku^3 的阻力，其中 u 为速度，k 为常数. 试用到初始位置的距离 s 和初速度 v_0 求出速度 (u) 和时间 (t) 的表达式.

2. 设一单位质量的质点沿 x 轴运动，并受到力 $f(x) = -\sin x$ 的作用.

(a) 如果在时刻 $t = 0$，质点在 $x = 0$，速度为 2，试确定该质点的运动. 证明当 $t \to \infty$ 时质点趋于极限位置. 试求这个极限位置.

(b) 如果除 v_0 可以有任意值之外其他条件都相同，试证如 $v_0 > 2$，那么当 $t \to \infty$ 时质点运动到无穷远；如果 $v_0 < 2$，那么质点在原点左右振动.

3. 选定原点在地球的中心的坐标轴，我们用 R 表示地球的半径. 按牛顿万有引力定律地球有 $-\mu M/y^2$ 的力吸引一个在 y 轴上的单位质量的质点，其中 μ 为"万有引力常数"，M 为地球的质量.

(a) 计算质点在点 $y_0(>R)$ 处放出以后的运动；即在时刻 $t = 0$ 时质点位于点 $y = y_0$ 且有速度 $v_0 = 0$ 这一条件下的运动.

(b) 试求在 (a) 中的质点击中地球时的速度.

(c) 利用 (b) 的结果，计算质点从无穷远落到地球时的速度[1].

*4. 受到轻微扰动的一质点在重力影响下从圆的顶点的静止位置滑下来. 这质点在哪一点不受约束地飞离这个圆?

*5. 质量为 m 的质点沿椭圆 $r = \dfrac{k}{1 - \varepsilon\cos\theta}$ 运动. 在质点上指向原点的力为 cm/r^2. 描述这个质点的运动，求出它的周期并且证明质点的向径在相等的时间内扫过相等的面积.

4A.1 节第 447 页

1. 证明外摆线（第 354 页例）的渐屈线是和第一个外摆线类似的外摆线，它可以从第一个外摆线经过旋转和收缩而得到.

2. 证明内摆线（第 355 页例）的渐屈线是另一个内摆线，它可以从第一个内摆线经过旋转和展开而得到.

---

1) 这与发射一个抛射体使它离开地球，而不再返回所必须的最小速度相同.

第五章　泰勒展开式

5.1　引言：幂级数

在微积分发展的早期，获得了一项巨大的成就，就是由牛顿和其他科学家发现了许多已知函数能够表示成“无穷次多项式”或“幂级数”，其系数是由一些极其优美而简明的规律形成的．例如 $\frac{1}{(1-x)}$ 或 $\frac{1}{(1+x^2)}$ 在开区间 $|x|<1$ 里成立的几何级数

$$\frac{1}{1-x}=1+x+x^2+\cdots+x^n+\cdots, \tag{1}$$

$$\frac{1}{1+x^2}=1-x^2+x^4-x^6+\cdots+(-1)^n x^{2n}+\cdots, \tag{1a}$$

就是它们的典型(参看第一章第 70 页)．

对许多其他的函数，带有数值系数 a_ν 的类似的展开式

$$f(x)=a_0+a_1x+\cdots+a_nx^n+\cdots$$
$$=\sum_{\nu=0}^{\infty}a_\nu x^\nu,$$

将在这一章里导出．

下面是一些值得注意的例子：

$$e^x=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots+\frac{x^n}{n!}+\cdots;$$

$$\sin x=x-\frac{x^3}{3!}+\frac{x^5}{5!}-\cdots+\cdots+\frac{(-1)^n x^{2n+1}}{(2n+1)!}+\cdots;$$

$$\cos x=1-\frac{x^2}{2!}+\frac{x^4}{4!}+\cdots+\frac{(-1)^n x^{2n}}{(2n)!}+\cdots$$

这些级数展开式对所有的 x 都成立．

牛顿的一般二项式定理． 展开式

$$(1+x)^{\alpha}=1+\frac{\alpha}{1!}x+\frac{\alpha(\alpha-1)}{2!}x^2+\cdots$$

$$=\sum_{\nu=0}^{\infty}\binom{\alpha}{\nu}x^{\nu}$$

对 $|x|<1$ 和任何指数 α 都是正确的.

为了说明这种展开式的确切意义，我们考虑由级数的前 $n+1$ 项的和所形成的 n 阶多项式，即第 n 个“部分和”

$$S_n=\sum_{\nu=0}^{n}a_{\nu}x^{\nu}.$$

公式

$$f(x)=\sum_{\nu=0}^{\infty}a_{\nu}x^{\nu}\quad(\text{式中}\quad |x|<a)$$

的意义就是当 $n\to\infty$ 时，在区间 $|x|<a$ 里，序列 S_n 在每一点 x 处都趋向于函数 $f(x)$ 的值. 这时，就说这个无穷级数在区间 $|x|<a$ 里收敛到 $f(x)$. 差

$$R_n(x)=f(x)-S_n(x)$$

称为级数的余项，它度量出在 x 这一点处以多项式 $S_n(x)$ 来逼近 $f(x)$ 的精确程度. 例如，

$$\frac{1}{1-x}=1+x+x^2+\cdots+x^n+R_n(x),\qquad(1b)$$

其中余项 $R_n(x)=x^{n+1}/(1-x)$，对 $|x|<1$，当 n 增加时趋于零. 这样就得到了无穷几何级数 $\sum_{\nu=0}^{\infty}x^{\nu}=1/(1-x)$. 要寻找在特殊情形下 R_n 的简单的容易处理的估计量，不论在理论上还是在实际上都是一个重要的任务.

在这一章里我们考虑广泛的一类函数的展开式，其中包括全部“初等”超越函数. 在这些超越函数的展开式里，一个很引人注意的事实是，系数是整数的优美的表达式. 这些展开式是通过泰勒定理得到的；以后在第七章，我们还将通过对幂级数的直接研究的另一途径来讨论它们.

应该强调指出的是，正如对于(1a)的几何级数一样，无穷展开式对于 x 在某个区间以外的值是不成立的（在几何级数的情况下这个区间是 $x^2 < 1$），即使由这个级数所表示的函数在这个区间以外有明确的定义，也还是这样.

5.2 对数和反正切的展开式

a. 对数函数

作为简单的例子，我们首先从具有余式 $r_n(t) = t^n/(1-t)$ 的几何级数

$$\frac{1}{1-t} = 1 + t + t^2 + \cdots + t^{n-1} + r_n(t)$$

出发，用积分来推导对数和反正切函数的展开式.

我们把这个和代入公式

$$-\log(1-x) = \int_0^x \frac{dt}{1-t}$$

中的被积函数，并且逐项积分，则对 $x < 1$ 得到

$$-\log(1-x) = x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots + \frac{x^n}{n} + R_n(x),$$

其中余项

$$R_n(x) = \int_0^x r_n dt = \int_0^x \frac{t^n}{1-t} dt.$$

因此对任何正整数 n，函数 $-\log(1-x)$ 都由 n 次多项式

$$x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots + \frac{x^n}{n}$$

近似地表示出来了，而余项 R_n 指出了这个近似式的“误差”.

为了评价这个近似式的精确度，我们来估计余项 R_n. 如果我们首先假定 $-1 \leqslant x \leqslant 0$，那么被积函数 $t^n/(1-t)$ 的绝对值在整个积分区间里到处都不会超过 $|t^n| = (-1)^n t^n$. 于是

$$|R_n| \leqslant \left|\int_0^x t^n dt\right| = \frac{|x|^{n+1}}{n+1};$$

因此对 x 在闭区间 $-1 \leqslant x \leqslant 0$ 里的每一个值，包括 $x = -1$，

选择足够大的 n，就能使这个余项变得像我们所希望的那样小(参看第 68 页)．对于 $x>0$，必须排除端点 $x=1$；我们不得不把 x 限制在半开区间 $0\leqslant x<1$ 里；被积函数不会改变符号，并且它的绝对值不会超过 $t^n/(1-t)$；于是对 $0\leqslant x<1$ 我们得到估计式

$$|R_n|\leqslant\frac{1}{1-x}\int_0^x t^n dt=\frac{x^{n+1}}{(1-x)(n+1)}.$$

因此我们又有，若令 x 固定，则当 n 充分大时，余项就可以任意小．当然这个估计对 $x=1$ 是没有意义的．

综上所述，只要 x 位于半开区间 $-1\leqslant x<1$ 里，都有

$$\log(1-x)=-x-\frac{x^2}{2}-\frac{x^3}{3}-\cdots-\frac{x^n}{n}-R_n,\quad(2)$$

其中余项 R_n 当 n 增大时，都趋于零．

事实上，这个推理建立了余项的一个"一致"的估计，即这一估计与 x 无关，对于 x 在区间 $-1\leqslant x\leqslant 1-h$ 里所有的值都成立，其中 h 为适合于 $0<h\leqslant 1$ 的任意数；也就是，$|R_n|\leqslant 1/[(n+1)h]$．

余式 R_n 在半开区间 $-1\leqslant x<1$ 里趋于零这个事实，就可说成在这个区间里对数函数可用一个无穷级数[1]

$$\log(1-x)=-x-\frac{x^2}{2}-\frac{x^3}{3}-\frac{x^4}{4}-\cdots\quad(3)$$

给出．如果我们在这个级数中代入特殊值 $x=-1$，就得到值得注意的公式

$$\log 2=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+-\cdots.\quad(4)$$

这个关系式的发现曾给微积分的早期开拓者以深刻的印象．

对于开区间 $-1<x<1$，我们仅需把(2)里的 x 换成 $-x$，

1) 我们把它留给读者作为一个练习，事实上对 $|x|>1$ 的所有 x 值，余项不仅不能趋于零，而且当 n 增加时 $|R_n|$ 趋向无穷大，所以对这种 x 的值，(2)中的多项式就不是这个对数函数的好的近似式，而且当 n 增加时反而变得更坏．

就得到

$$\log(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+-\cdots+(-1)^{n-1}\frac{x^n}{n}-R'_n,\quad (2a)$$

其中

$$R'_n(x)=\int_0^{-x}\frac{t^n dt}{1-t}=(-1)^{n+1}\int_0^x\frac{t^n dt}{1+t}.$$

取 n 为偶数，由 (2a) 减去 (2)，就有

$$\frac{1}{2}\log\left(\frac{1+x}{1-x}\right)=\text{ar tanh}\, x$$

$$=x+\frac{x^3}{3}+\frac{x^5}{5}+\cdots+\frac{x^{n-1}}{n-1}+\bar{R}_n,$$

其中余项 $\bar{R}_n$ 由

$$\bar{R}_n=\frac{1}{2}(R_n-R'_n)=\int_0^x\frac{t^n}{1-t^2}\,dt$$

给出，这里的 ar tanh$x$ 依第 247 页上的定义.

注意到 $1/(1-t^2)\leqslant 1/(1-x^2)$，由这个积分的一个基本估计式，我们得到

$$|\bar{R}_n|\leqslant\frac{|x^{n+1}|}{n+1}\cdot\frac{1}{1-x^2};$$

所以，当 n 增加时余项 $\bar{R}_n$ 趋于零. 这里我们又得到一个展成无穷级数的例子：

$$\frac{1}{2}\log\left(\frac{1+x}{1-x}\right)=\text{ar tanh}\, x=x+\frac{x^3}{3}+\frac{x^5}{5}+\cdots,\quad (5)$$

对 $|x|<1$ 的所有 x 的值都成立. 顺便说明一下，此结果也能由直接积分 $1/(1-x^2)$ 的几何级数而得到. 这个公式的一个优点是，当 x 从 -1 到 $+1$ 变化时，表达式 $(1+x)/(1-x)$ 随之而取遍全部正数. 因此，如果把 x 选得适当，我们就可以通过这个级数去计算任何正数的对数值，其误差不会超过以上的估计量 $\bar{R}_n$.

b. 反正切函数

我们能够用和对数函数有些类似的方法，从公式

$$\frac{1}{1+t^2}=1-t^2+t^4-+\cdots+(-1)^{n-1}t^{2n-2}+r_n$$

出发来处理反正切函数，其中 $r_n=(-1)^n\dfrac{t^{2n}}{1+t^2}$.

通过积分(参看第 228 页)，我们得到

$$\arctan x=x-\frac{x^3}{3}+\frac{x^5}{5}-+\cdots+(-1)^{n-1}\frac{x^{2n-1}}{2n-1}+R_n,$$

$$R_n=(-1)^n\int_0^x\frac{t^{2n}}{1+t^2}\,dt;$$

我们立即看到余项 R_n 在闭区间 $-1\leqslant x\leqslant 1$ 里随 n 的增加而趋于零，因为

$$|R_n|\leqslant\int_0^{|x|}t^{2n}\,dt=\frac{|x|^{2n+1}}{2n+1}.$$

从余项的公式，我们也能很容易地证明，在 $|x|>1$ 时，余项的绝对值随 n 的增加而无限增大.

因此，我们就推导出无穷级数

$$\arctan x=x-\frac{x^3}{3}+\frac{x^5}{5}-+\cdots+(-1)^{n-1}\frac{x^{2n-1}}{2n-1}+-\cdots, \tag{6}$$

对区间 $|x|\leqslant 1$ 成立. 因为对 $x=1$ 有 $\arctan 1=\dfrac{\pi}{4}$，我们得到莱布尼兹-格雷戈里级数 (Leibnitz-Gregory):

$$\frac{\pi}{4}=1-\frac{1}{3}+\frac{1}{5}-+\cdots, \tag{7}$$

这个表达式和前面建立的 log 2 的表达式一样值得注意.

5.3 泰勒定理

牛顿的弟子泰勒 (Tayler) 注意到多项式的初等展开可以广泛地推广到非多项式函数，只要这些函数充分可微，而区域又受到适当的限制.

a. 多项式的泰勒表示

关于 x 的一个 n 次多项式，譬如说

$$f(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n,$$

其泰勒表示完全是一个初等的代数式.

假如我们把 x 换成 $a+h=b$，并按 h 的幂展开每一项，就马上得到一个形式为

$$f(a+h)=c_0+c_1h+c_2h^2+\cdots+c_nh^n \tag{8}$$

的表达式. 泰勒公式实质上就是关于这些系数的关系式

$$c_\nu=\frac{1}{\nu!}f^{(\nu)}(a), \tag{8a}$$

系数 c_ν 由 f 和 f 的各阶导数在 $x=a$ 的值来表示. 为了证明这个事实，我们把量 $h=b-a$ 看作自变量，并应用锁链法则；根据这个法则关于 h 求导和关于 $b=a+h$ 求导结果是一样的. 因此在公式(8)里关于 h 逐次求导，在每次求导之后以 $h=0$ 代入，就依次地得到如下的结果：

$$c_0=f(a),\quad c_1=f'(a),\cdots,\quad \nu!c_\nu=f^{(\nu)}(a),$$

因而的确得到了多项式的泰勒公式：

$$f(a+h)=f(a)+hf'(a)+\frac{h^2}{2!}f''(a)+\cdots+\frac{h^n}{n!}f^{(n)}(a). \tag{9}$$

因为 n 次多项式的 $n+1$ 次导数是零，所以公式 (9) 就自然而然地结束了.

如上所述，公式 (9) 不过是把一个关于 $a+h$ 幂的多项式，经初等代数整理而化为关于 h 幂的一个多项式.

b. 非多项式函数的泰勒公式

牛顿和他最亲近的弟子勇敢地把公式 (9) 用到了非多项式函数. 对这种非多项式函数来说，展开式并不自动地在第 n 项结束；他们简单地代之以允许 n 趋向于无穷，以后将验证这种处理办法对许多重要的特殊函数是合理的.

假设函数 f 在包含 a 和 $a+h$ 的一个区间里至少 n 次可微，我们当然不能再把 $f(a+h)$ 写成象公式(9)那样用 h 的有穷次幂构成的表达式，而必须通过一个附加的"余项" R_n 进行修正. 我们暂且试探性地把它写成

$$f(b)=f(a+h)$$
$$=f(a)+hf'(a)+\cdots+\frac{h^n}{n!}f^{(n)}(a)+R_n; \quad (10)$$

事实上，(10) 只不过是校正式余项 R_n 的一个定义，并且指出当 $n\to\infty$ 时，希望 R_n 变小而趋于零. 如果这个余项确实趋于零，那么公式 (10) 在 $n\to\infty$ 时的极限就导出了 $f(x)$ 的一个表成 h 的无穷幂级数的展开式

$$f(a+h)=f(a)+hf'(a)+\cdots+\frac{h^n}{n!}f^{(n)}(a)+\cdots. \quad (11)$$

于是关键性的问题在于寻求余项 R_n 的估计式，以便能够严格地揭示出关于 h 的 n 阶有限泰勒多项式

$$T_n(h)=\sum_0^n\frac{f^{(\nu)}(a)}{\nu!}h^\nu \quad (12)$$

作为泰勒表示的精确度，以及当 $n\to\infty$ 时取极限的合理性；这个问题的难度远远超过了在5.3a这一节里代数处理的难度. 泰勒多项式 $T_n(h)$ 是 $f(a+h)$ 在这样意义下的一个近似式，即在 $h=0$ 处不但函数 T_n 和 f 而且它们的直到 n 阶的导数都相重合，使得差 $R_n=f-T_n$ 连同它们的前 n 阶导数在 $x=a$ 一起变为零.

5.4 余项的表示式及其估计

a. 柯西和拉格朗日余项

余项 R_n 的一个直接的表示，使得我们有可能对它的绝对值 $|R_n|$ 作出估计，这是泰勒定理的核心. 这些结果容易根据微分中值定理而得到. 此外它们还与用微分作函数的线性近似有关 (参看第 190 页).

首先让我们再一次考察一下这个近似式.

在点 a 处的导数的定义仅仅是说 $f(a+h)=f(a)+hf'(a)+h\varepsilon$，这里当 $h\to 0$ 时 $\varepsilon\to 0$. 只要假定不仅 f' 而且 f'' 在区间 J 里存在并且是连续的，我们就能够肯定 ε 确实至少和 h 同阶，从而得到稍微精确一些的近似表达式. 如果我们再一次记 $a+h=b$，由

$$f(b)=f(a)+(b-a)f'(a)+R \tag{13}$$

引入一个余项 R，我们就可以得到上述关于 ε 的估计. 现在把 b 看成是固定的，把初始点 a 看成是变量；方程(13)在 J 中把 R 定义为 a 的函数，对上式两边取 a 的导数，因为左边的 $f(b)$ 是常量，故导数为零，对右边应用两个函数乘积的微分法则，就证明了

$$0=f'(a)-f'(a)+(b-a)f''(a)+R'(a),$$

从而得到

$$-R'(a)=(b-a)f''(a). \tag{14}$$

现在，对 $a=b$ 我们显然有 $R(b)=0$. 由微商中值定理 $[R(a)-R(b)]/(b-a)=-R'(\xi)$，其中 ξ 是未被指明的介于 a 和 b 中间的一个值；由于 $R(b)=0$，我们由此断定 $R(a)=-(b-a)\times R'(\xi)=-hR'(\xi)$. 现在由(14)有 $R'(\xi)=-(b-\xi)f''(\xi)$，而由于 $|b-\xi|\leqslant h$，所以 $|R'(\xi)|<h|f''(\xi)|$. 因为 $|f''(\xi)|$ 在围绕 a 的一个区间里是有界的，最后我们得到一个估计式，它指明余项或“误差” R_n 关于 h 至少是二阶无穷小：

$$|R(a)|<h^2|f''(\xi)|. \tag{15}$$

我们现在从 $n=1$ 的特殊情形转向任意阶 n 的情形. 可直接用与 $n=1$ 的情形相同的方法来描述余项 R_n. 假定 a 和 $b=a+h$ 是区间 J 里的点，在这个区间中 $f(x)$ 有定义并直到 $n+1$ 阶连续可微. 把 a 看作自变量，而终点 b 保持固定，在第 470 页定义 $R_n(a)$ 的公式(10)里，我们用 $b-a$ 来代替 h. 求导数，并考虑到 $f(b)$ 是常数，根据乘积的微分规则我们就发现，几乎所有的项都抵消了，剩下的只是公式：

$$0=\frac{(b-a)^n}{n!}f^{(n+1)}(a)+R_n'(a), \tag{16}$$

对区间里的每个值 a 都成立. 因为对于 $a=b$, 余项 R_n 是零,所以这个把 R_n 的导数直接表为 a 的函数的表达式完全刻划了 R_n: 把它表成了积分 $\int_b^a R_n'(t)\,dt=-\int_a^b R_n'(t)\,dt$, 也就是

$$R_n(a)=\int_a^b \frac{(b-t)^n}{n!} f^{(n+1)}(t)\,dt. \tag{17}$$

这是余项的一个精确的积分表示式.

余项 R_n 的一个类似于上述 $n=1$ 时所得到的估计, 可以对通过公式 (16) 应用微分中值定理而直接得到:

$$\frac{R_n(a)-R_n(b)}{b-a}=\frac{R_n(a)}{b-a}=-R_n'(\xi)=\frac{(b-\xi)^n}{n!}f^{(n+1)}(\xi)$$

或

$$R_n(a)=\frac{(b-a)(b-\xi)^n}{n!} f^{(n+1)}(\xi), \tag{18}$$

其中 ξ 是介于 a 和 b 之间而未被指明的一个适当的中间值. 对表达式 (17) 应用积分中值定理, 也能得到相同的估计式(第二章第 151 页).

余项的柯西 (Cauchy) 形式. 如果我们定义 $\xi=a+\theta h=a+\theta(b-a)$,就得到泰勒公式(10)里的余项的柯西形式

$$R_n(a)=\frac{h^{n+1}}{n!}(1-\theta)^n f^{(n+1)}(a+\theta h), \tag{19}$$

其中 θ 是 0 和 1 之间的一个未指明的量.

对余项 R_n 的积分形式 (17), 我们也能用积分学的推广了的中值定理(见第 152 页),取表达式 $p(t)=(b-t)^n$ 作为“权函数”; 这个函数在整个积分区间上都不变号.[1] 这样就有

$$R_n=\frac{1}{n!} f^{(n+1)}(\xi)\int_a^b (b-t)^n dt=\frac{(b-a)^{n+1}}{(n+1)!}f^{(n+1)}(\xi). \tag{20}$$

余项的拉格朗日 (Lagrange) 形式. 再次令 $\xi=a+\theta h$ 就得

1) 对于 $p(t)$ 为正的情形广义中值定理已证明了, 而当 $p(t)$ 在整个积分区间是负的时候,它也完全同样适用.

到余项的拉格朗日形式

$$R_n(a)=\frac{h^{n+1}}{(n+1)!}f^{(n+1)}(a+\theta h),\qquad (21)$$

其中 θ 是满足 $0\leqslant\theta\leqslant 1$ 的一个适当的数. 拉格朗日形式特别具有启发性,因而更经常地被应用,因为它使得公式

$$f(a+h)=f(a)+\frac{h}{1!}f'(a)+\frac{h^2}{2!}f''(a)+\cdots$$

$$+\frac{h^n}{n!}f^{(n)}(a)+R_n=P_n(h)+R_n\qquad (22)$$

的余项 R_n, 看起来很像是在展开式 (22) 里以更高一阶出现的项 $h^{n+1}f^{(n+1)}(a)/(n+1)!$, 只是自变量 a 被中间值 $a+\theta h$ 代替罢了.

若函数 f 的 $n+1$ 阶导数 $f^{(n+1)}$ 在包含点 a 的一个闭区间里是连续的,则量 $|f^{(n+1)}(\xi)|$ 就有一个固定的界 M. 于是因为

$$|R_n|\leqslant\frac{h^{n+1}}{(n+1)!}M,$$

所以对于固定的 n 泰勒多项式 $P_n(h)$ 给出了函数 $f(a+h)$ 的一个近似式,其误差关于 h 的阶至少是 $n+1$.

我们的兴趣主要将针对这样的问题,当 n 增加时余项 R_n 是否趋于零;如果趋向于零,我们就说函数展成了一个无穷的泰勒级数

$$f(a+h)=f(a)+\frac{h}{1!}f'(a)+\frac{h^2}{2!}f''(a)$$

$$+\frac{h^3}{3!}f'''(a)+\cdots;\qquad (23)$$

特别是,若先令 $a=0$, 而后把 h 改写成 x, 我们就得到"幂级数"

$$f(x)=f(0)+\frac{x}{1!}f'(0)+\frac{x^2}{2!}f''(0)+\cdots.$$

我们将在第 5.5 节里讨论一些例子.

这个对于固定的 n 的、有限的、并带有余项的泰勒展开式(22), 在应用上同样是重要的. 在这个公式里,如果我们令 h 趋于零,并

用第三章第 268 页上的术语，我们就可以说，级数的各项就都以关于 h 的不同的阶趋于零. 在泰勒级数里，表达式 $f(a)$ 表示零阶的项，表达式 $hf'(a)$ 表示第一阶的项，表达式 $h^2f''(a)/2!$ 表示二阶的项，等等. 我们从余项的形式看到，在展开一个函数直到第 n 阶项时，就形成一个误差，当 h 趋于零时，这个误差趋于零的阶数为 $n+1$. 点 $a+h$ 越接近于点 a，函数 $f(a+h)$ 用近似多项式 $P_n(h)$ 来表示就越好；在一些最重要的情况下，在 x 的一个邻域里，这个近似表达式能够由增大 n 值而得到改进.

b. 泰勒公式的另一种推导法

泰勒定理中余项 R_n 的积分表示式 (17) 是以 R_n' 的公式 (16) 为基础的. 由于这个定理的重要性，我们在这里从另一角度再给出一种推导法：由公式

$$f(b)-f(a)=\int_a^b f'(t)\,dt \tag{24}$$

出发，经反复运用分部积分法而直接导出余项 R_n 的表达式.

为了通过逐次分部积分来变换 (24)，我们按照关系式

$$\phi_0(t)=1, \quad \phi_\nu'(t)=\phi_{\nu-1}(t) \tag{25}$$

和条件

$$\phi_\nu(b)=0 \quad (\nu\geqslant 1) \tag{26}$$

来引进函数

$$\phi_1(t),\ \phi_2(t),\ \cdots,\phi_\nu(t),\ \cdots,$$

其中 b 作为固定的参数. 很清楚，条件 (25) 和 (26) 逐步地确定了所有的 $\phi_\nu(t)$. 容易直接地验证，$\phi_\nu(t)$ 刚好是多项式：

$$\phi_\nu(t)=\frac{(t-b)^\nu}{\nu!}.$$

我们顺便注意一下，函数 ϕ_ν 是通过反复积分逐个产生的，积分常数留下待定；所以定义条件 (25) 也能够被用满足另外一组边条件 (26) 的函数所满足(见第 199 页).

因为 $\phi_\nu(a)=(-1)^\nu(b-a)^\nu/\nu!$ 而 $\phi_\nu(b)=0$，我们得到

$$f(b)-f(a)=\int_a^b \phi_0 f'\,dt=\int_a^b \phi_1' f'\,dt=\phi_1 f'\Big|_a^b-\int_a^b \phi_1 f''\,dt;$$

对最后一项再次进行分部积分，就得到

$$f(b)-f(a)=(b-a)f'(a)-\int_a^b \phi_2' f'' dt$$

$$=(b-a)f'(a)+\frac{(b-a)^2}{2!}f''(a)+\int_a^b \phi_2 f''' dt;$$

把这个过程反复进行 n 次就有

$$f(b)-f(a)=(b-a)f'(a)+\frac{(b-a)^2}{2!}f''(a)+\cdots$$

$$+\frac{(b-a)^n}{n!}f^{(n)}(a)+\int_a^b(-1)^n\phi_n(t)f^{(n+1)}(t)dt$$

$$=(b-a)f'(a)+\frac{(b-a)^2}{2!}f''(a)+\cdots$$

$$+\frac{(b-a)^n}{n!}f^{(n)}(a)+R_n,$$

其中按照 ϕ_n 的定义，

$$R_n=\int_a^b f^{(n+1)}(t)\frac{(b-t)^n}{n!}dt.$$

这样我们就再一次证明了：

泰勒定理. 如果一个函数 $f(t)$ 在包含点 a 和 b 的一个闭区间上直到第 $n+1$ 阶连续可微，那末：

$$f(b)=f(a)+(b-a)f'(a)+\cdots+\frac{(b-a)^n}{n!}f^{(n)}(a)+R_n,$$

其中余式 R_n 依赖于 n，a 和 b，并由表达式

$$R_n=\frac{1}{n!}\int_a^b(b-t)^n f^{(n+1)}(t)dt \tag{27}$$

给出.

把记号改一改，我们就得到泰勒公式的另一个稍微不同的表达式. 把 a 换成 x，b 换成 $x+h$，我们就有

$$f(x+h)=f(x)+hf'(x)+\cdots+\frac{h^n}{n!}f^{(n)}(x)+R_n, \tag{27a}$$

其中

$$R_n = \frac{1}{n!}\int_x^{x+h}(x+h-t)^n f^{(n+1)}(t)dt$$

或用 $t = x + \tau$,

$$R_n = \frac{1}{n!}\int_0^h (h-\tau)^n f^{(n+1)}(x+\tau)d\tau. \qquad (27b)$$

如果我们令 $x = 0$ 并且把 h 改写成 x，我们就得到[1]

$$f(x) = f(0) + \frac{x}{1!}f'(0) + \frac{x^2}{2!}f''(0) + \cdots + \frac{x^n}{n!}f^{(n)}(0) + R_n, \qquad (27c)$$

其中余式为

$$R_n = \frac{1}{n!}\int_0^x (x-t)^n f^{(n+1)}(t)dt.$$

对这个积分应用积分学中值定理或它的推广的形式，就分别导出余项的柯西公式

$$R_n = \frac{(1-\theta)^n}{n!} x^{n+1} f^{(n+1)}(\theta x)$$

和拉格朗日公式

$$R_n = \frac{x^{n+1}}{(n+1)!} f^{(n+1)}(\theta x),$$

同以前(第 472 页)证明了的一样.

这里 θ 是一个满足 $0 \leqslant \theta \leqslant 1$ 的适当的未指明的数(两个公式中的 θ 并不相同).

作为一个练习，读者应自己构造 ϕ_ν，使它满足(25)，而边条件(26)将代以关系式

$$\int_0^1 \phi_\nu(t)dt = 0 \qquad (\nu \geqslant 1)$$

(见第八章，附录 A).

---

1) 定理的这个特殊情形有时称为马克劳林 (Maclaurin) 定理是不符合历史的；泰勒的一般定理在 1715 年已公开发表；而马克劳林的特殊结果却在 1742 年.

5.5 初等函数的展开式

前面所得的一般结论容许我们将简单的初等函数展开成泰勒级数．其他函数的展开式将在第七章中讨论．

a. 指数函数

首先我们展开指数函数 $f(x)=e^x$．在这个情形下，所有的导数都和 $f(x)$ 相同，并且对 $x=0$ 其值都是 1．根据拉格朗日的余项形式(第 473 页，等式(21))，立即得到公式：

$$e^x = 1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots+\frac{x^n}{n!}+\frac{x^{n+1}}{(n+1)!}e^{\theta x},$$
$$0<\theta<1.$$

如果我们现在令 n 无限增大，那末对于任何一个固定的 x 值，余式 R_n 都趋于零．为了证明这一点我们首先注意到 $e^{\theta x}\leqslant e^{|x|}$ (因为 e^x 是单调增加的函数)．令 m 是大于 $2|x|$ 的任一整数．则对所有的 $k\geqslant m$ 都有 $|x|/k<\frac{1}{2}$，因而

$$\left|\frac{x^{n+1}}{(n+1)!}\right| = \frac{|x^m|}{m!}\cdot\frac{|x|}{m+1}\cdots\frac{|x|}{n+1}$$
$$\leqslant \frac{|x^m|}{m!}\cdot\frac{1}{2^{n+1-m}}\leqslant\frac{|2x|^m}{m!}\cdot\frac{1}{2^n};$$

所以

$$|R_n|\leqslant\frac{|2x|^m}{m!}\cdot e^{|x|}\cdot\frac{1}{2^n}.$$

因为右边的前两个因子与 n 无关，而当 $n\to\infty$ 时 $1/2^n\to 0$，我们的命题就被证明了．

于是函数 e^x 就可以表成一个无穷级数：

$$e^x = 1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots$$
$$=\sum_{\nu=0}^{\infty}\frac{x^\nu}{\nu!}.$$

这个展开式对所有的 x 都正确．特别是对 $x=1$ 我们又得到第一章里用以定义数 e 的那个无穷级数(参看第 81 页)．

当然，对于数值计算，我们必须用带有余项的泰勒定理；例如以 $x=1$ 作为例子(与第 82 页上的相似的计算比较)我们有

$$e=1+1+\frac{1}{2!}+\frac{1}{3!}+\cdots+\frac{1}{n!}+\frac{e^{\theta}}{(n+1)!}.$$

如果我们想要计算 e 而要求误差至多为 1/10,000，我们仅需要选择足够大的 n，使得余项小于 1/10,000，因为这个余项必定小于 $3/(n+1)!$[1]，选择 $n=7$ 就够了，因为 $8!>30,000$．这样我们就得到 e 的近似值 $e=2.71825$，其误差小于 0.0001．

b. $\sin x$，$\cos x$，$\sinh x$，$\cosh x$ 的展开式

对于函数 $\sin x$，$\cos x$，$\sinh x$，$\cosh x$ 我们得到以下的公式：

$$\begin{array}{lllll}
f(x)= & \sin x & \cos x & \sinh x & \cosh x, \\
f'(x)= & \cos x & -\sin x & \cosh x & \sinh x, \\
f''(x)= & -\sin x & -\cos x & \sinh x & \cosh x, \\
f'''(x)= & -\cos x & \sin x & \cosh x & \sinh x, \\
f^{(4)}(x)= & \sin x & \cos x & \sinh x & \cosh x.
\end{array}$$

可见，对 $\sin x$ 和 $\sinh x$ 的关于 x 的近似多项式，x 的偶次幂的系数为零，而对 $\cos x$ 和 $\cosh x$，x 的奇次幂的系数为零．

当我们用余项的拉格朗日形式(21)时(第 473 页)，上述函数的泰勒级数取以下的形式：

$$\sin x=x-\frac{x^3}{3!}+\frac{x^5}{5!}+\cdots+\frac{(-1)^n x^{2n+1}}{(2n+1)!}$$
$$+\frac{(-1)^{n+1}x^{2n+3}\cos(\theta x)}{(2n+3)!},$$

1) 这里我们利用了 $e<3$ 这个事实．这可以从 e 的级数得出(参看第81页)，因为 $1/n!\leqslant 1/2^{n-1}$ 总是正确的，所以

$$e<1+1+\frac{1}{2}+\frac{1}{4}+\cdots=1+\frac{1}{1-\frac{1}{2}}=3.$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + \cdots + \frac{(-1)^n x^{2n}}{(2n)!} + \frac{(-1)^{n+1} x^{2n+2} \cos(\theta x)}{(2n+2)!},$$

$$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + \frac{x^{2n+1}}{(2n+1)!} + \frac{x^{2n+3}\cosh(\theta x)}{(2n+3)!},$$

$$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + \frac{x^{2n}}{(2n)!} + \frac{x^{2n+2}\cosh(\theta x)}{(2n+2)!}.$$

当然，在四个公式里的每一个当中，θ 表示区间 $0 \leqslant \theta \leqslant 1$ 上不同的数，此外 θ 还依赖于 n 和 x。因为在这些公式中的每一个里当 n 增大时余项都趋于零，这恰恰和 e^x 所看到的情形完全相同，所以我们就能够使近似要怎样精确就可以怎样精确。这样我们又得到四个无穷级数，它们对所有的 x 都成立：

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + \cdots = \sum_{\nu=0}^{\infty} \frac{(-1)^\nu x^{2\nu+1}}{(2\nu+1)!},$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + \cdots = \sum_{\nu=0}^{\infty} \frac{(-1)^\nu x^{2\nu}}{(2\nu)!},$$

$$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots = \sum_{\nu=0}^{\infty} \frac{x^{2\nu+1}}{(2\nu+1)!},$$

$$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots = \sum_{\nu=0}^{\infty} \frac{x^{2\nu}}{(2\nu)!}.$$

后两个也可以根据双曲函数的定义(见第245页)从 e^x 的级数展开式得到.

c. 二项式级数

我们不讨论已经在第 5.2 节直接推导过的函数 $\log(1+x)$ 和

$\arctan x$ 的泰勒级数了；但是我们将着手处理任意指数的广义的二项式定理，这是牛顿的数学发现中最引人注意的一个．我们希望展开函数 $f(x)=(1+x)^{\alpha}$ 成泰勒级数，其中 $x>-1$，而 α 是一个任意的数，正的或负的，有理的或无理的．我们选择函数 $(1+x)^{\alpha}$ 以代替 x^{α}，因为对后者来说在 $x=0$ 并非所有的导数都是连续的，除非 α 为正整数值的这种明显的情形．首先我们计算 $f(x)$的导数，得到

$$f'(x)=\alpha(1+x)^{\alpha-1},$$
$$f''(x)=\alpha(\alpha-1)(1+x)^{\alpha-2},\cdots,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$f^{(\nu)}(x)=\alpha(\alpha-1)\cdots(\alpha-\nu+1)(1+x)^{\alpha-\nu}.$$

特别是，对 $x=0$ 我们有

$$f'(0)=\alpha,\qquad f''(0)=\alpha(\alpha-1),\cdots,$$
$$f^{(\nu)}(0)=\alpha(\alpha-1)\cdots(\alpha-\nu+1).$$

于是泰勒定理就写成为

$$(1+x)^{\alpha}=1+\alpha x+\frac{\alpha(\alpha-1)}{2!}x^2+\cdots$$
$$+\frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)}{n!}x^n+R_n.$$

收敛性

我们还必须讨论余项．这个问题并不很困难，但也并不象前面已探讨过的那样简单．我们一方面将直接得到余项的一个估计，另方面也将作为 A.4 节里一般结果的特殊情形．这将允许我们断定 $|x|<1$ 时二项展开式的余项 R_n 趋于零．因此表达式 $(1+x)^{\alpha}$ 可以展成一个无穷级数

$$(1+x)^{\alpha}=1+\frac{\alpha}{1!}x+\frac{\alpha(\alpha-1)}{2!}x^2+\cdots$$
$$=\sum_{\nu=0}^{\infty}\binom{\alpha}{\nu}x^{\nu},$$

其中为了简洁，我们引进了广义二项系数

$$\binom{\alpha}{\nu} = \frac{\alpha(\alpha-1)\cdots(\alpha-\nu+1)}{\nu!} \quad (\nu > 0)$$

$$\binom{\alpha}{0} = 1.$$

*为了直接证明在 $-1 < x < 1$ 的情形下当 $n \to \infty$ 时余项 $R_n \to 0$，我们选用余项的柯西形式(19)(第 472 页)：

$$R_n = \frac{(1-\theta)^n}{n!} x^{n+1} f^{(n+1)}(\theta x)$$

$$= \frac{(1-\theta)^n}{n!} \alpha(\alpha-1)(\alpha-2)\cdots$$

$$\cdots(\alpha-n)x^{n+1}(1+\theta x)^{\alpha-n-1}$$

$(0 \leqslant \theta \leqslant 1)$. 因为 $|x| < 1$，我们有 $0 \leqslant (1-\theta)/(1+\theta x) \leqslant 1$，所以

$$|R_n| \leqslant (1+\theta x)^{\alpha-1}|\alpha x| \left|\left(1-\frac{\alpha}{1}\right)x\right| \left|\left(1-\frac{\alpha}{2}\right)x\right| \cdots$$

$$\cdots \left|\left(1-\frac{\alpha}{n}\right)x\right|.$$

存在一个数 q 满足 $|x| < q < 1$. 那么显然也满足

$$\left|\left(1-\frac{\alpha}{m}\right)x\right| < q,$$

只要 m 足够大，譬如说 $m > N$. 这样，对 $n > N$ 就有

$$|R_n| \leqslant (1+\theta x)^{\alpha-1}|\alpha|(1+|\alpha|)^N q^{n-N},$$

其中因子 $(1+x\theta)^{\alpha-1}$ 是有界的(如果 $\alpha \geqslant 1$，以 $2^{\alpha-1}$ 为界；如果 $\alpha < 1$，以 $(1-q)^{\alpha-1}$ 为界)，所以 $R_n \to 0$.

对表达式 $(a+b)^\alpha$ 可给出一个稍微广泛些的公式. 我们只要提出因子 a^α 并且对 $x = b/a$ 应用二项展开式，则对 $a > 0$ 和 $|b| < a$ 得到

$$(a+b)^\alpha = a^\alpha\left(1+\frac{b}{a}\right)^\alpha$$

$$= a^{\alpha}\left(1 + \alpha\frac{b}{a} + \frac{\alpha(\alpha-1)}{1\cdot 2}\left(\frac{b}{a}\right)^{c} + \cdots\right)$$

$$= a^{\alpha} + \frac{\alpha}{1}a^{\alpha-1}b + \frac{\alpha(\alpha-1)}{1\cdot 2}a^{\alpha-2}b^{2} + \cdots.$$

5.6 几 何 应 用

一个函数 $f(x)$ 在 $x=a$ 的一个邻域里的性状，或者说给定的曲线在它的一个点的邻域里的性状，可以借助于泰勒定理详细地加以刻划，因为这个定理允许我们对邻近的点 $x=a+h$ 把函数的增量分解成 h 的一阶，二阶等量的和.

a. 曲线的接触

高阶接触

如果在 $x=a$ 点，两条曲线 $y=f(x)$ 和 $y=g(x)$ 相交且有公共的切线，我们就说这两条曲线相互接触或有第一阶接触．在这种情形下，函数 $f(a+h)$ 和 $g(a+h)$ 的泰勒展开式关于 h 具有相同的零阶项和一阶项．若 $f(x)$ 和 $g(x)$ 的二阶导数在 $x=a$ 点也彼此相等，我们说这两条曲线有二阶接触．这时，f 和 g 的泰勒展开式二阶的项也将一致．如果我们假定两函数至少三阶可微，那么差

$$D(x)=f(x)-g(x)$$

能够表示成以下的形式：

$$D(a+h)=f(a+h)-g(a+h)$$
$$=\frac{h^3}{3!}D'''(a+\theta h)=\frac{h^3}{3}F(h),$$

其中表达式 $F(h)$ 当 h 趋向于零时趋于 $f'''(a)-g'''(a)$．因此，差 $D(a+h)$ 趋于零关于 h 至少是三阶的.

我们能够用这种方法来讨论一般的情形，这时 $f(x)$ 和 $g(x)$ 的泰勒级数直到第 n 阶项都一致；即，

$$f(a)=g(a),\quad f'(a)=g'(a),\cdots,f^{(n)}(a)=g^{(n)}(a).$$

我们假定第 $n+1$ 阶导数是连续的. 在这些条件下，由我们的两个函数定义的曲线就说是在 $x=a$ 点有 n 阶接触. 这时这两个函数的差有以下的形式:

$$\begin{aligned} D(a+h) &= f(a+h)-g(a+h) \\ &= \frac{h^{n+1}}{(n+1)!} D^{(n+1)}(a+\theta h) \\ &= \frac{h^{n+1}}{(n+1)!} F(h), \end{aligned}$$

其中因为 $0 \leqslant \theta \leqslant 1$，量 $F(h)=D^{(n+1)}(a+\theta h)$ 当 h 趋于零时趋于 $f^{(n+1)}(a)-g^{(n+1)}(a)$. 从这个公式我们看到，在此接触点上差 $f(x)-g(x)$ 趋于零至少是 $n+1$ 阶.

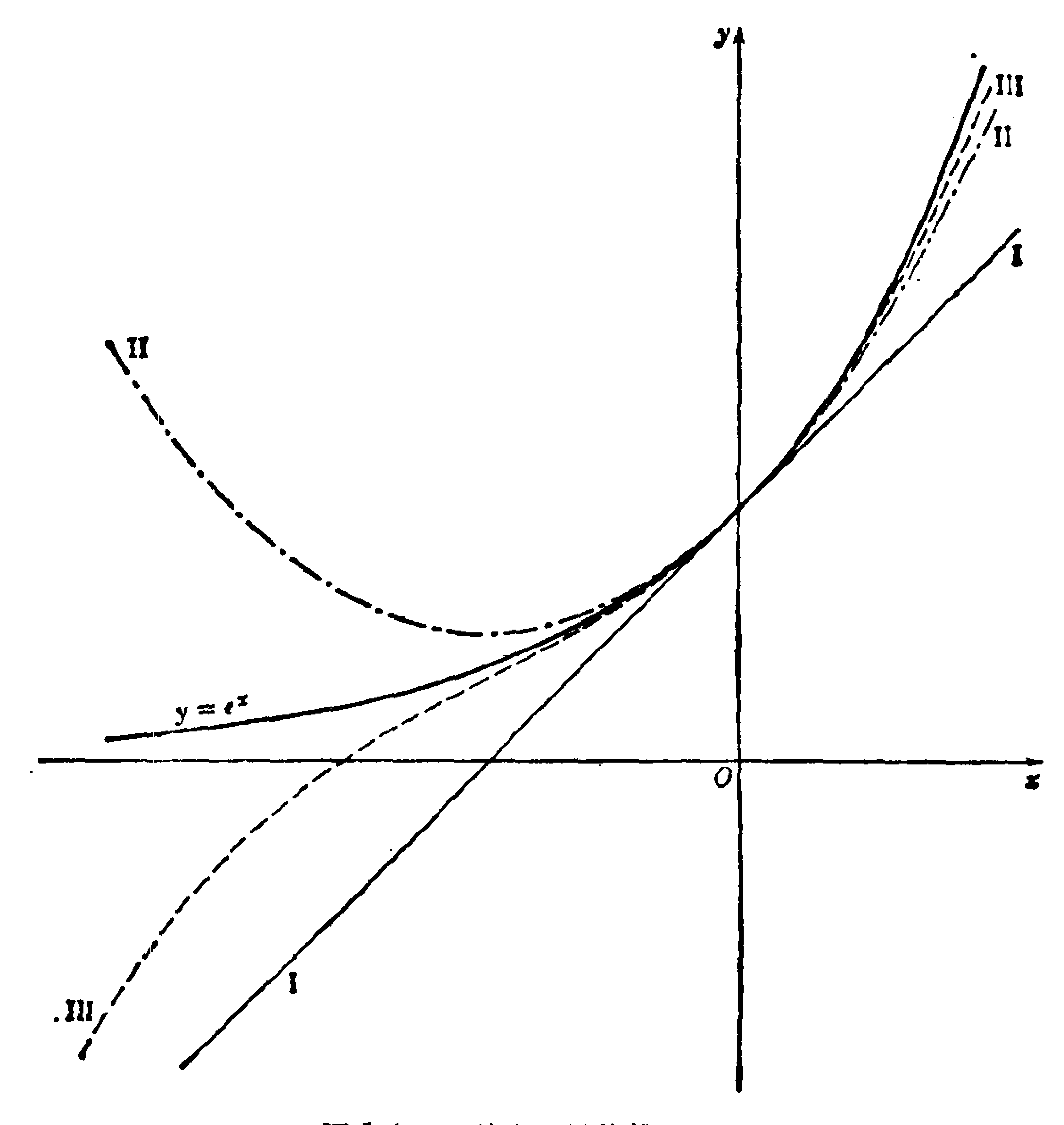


图 5.1 e^x 的密切抛物线

由等式

$$P_n(x)=f(a)+\frac{x-a}{1!}f'(a)+\cdots+\frac{(x-a)^n}{n!}f^{(n)}(a)$$

定义的泰勒多项式在几何上由这样一个 n 阶“抛物线”来刻划，它在给定的点上对给定函数的图形具有最大可能的接触。因此这些抛物线有时称为密切(osculating)抛物线(仅当 $n=2$ 时这些曲线才是普通意义下的抛物线)。

图 5.1 显示出函数 $y=e^x$ 在 $x=0$ 点的最初三条密切抛物线。

在 $x=a$ 点有 n 阶接触的两条曲线 $y=f(x)$ 和 $y=g(x)$，也可能有更高阶的接触，也就是说，等式 $f^{(n+1)}(a)=g^{(n+1)}(a)$ 可能也是正确的。如果不是这种情形，若 $f^{(n+1)}(a)\neq g^{(n+1)}(a)$，我们就说接触的阶恰好是 n[1]。

偶阶接触或奇阶接触

由公式同样也由于直观，我们指出一个值得注意的事实，它常常不为初学者所注意。设两条曲线的接触恰好是偶阶；也就是说，所讨论的两个函数的开头的某 n 个(偶数)导数都在所考虑的点上具有相同的值，而第 $n+1$ 阶导数则不同。那末对于 h 的小的正值和负值，上述公式表明 $f(a+h)-g(a+h)$ 这个差有不同的符号。于是这两条曲线在接触点上相交叉。例如，若第三阶导数有不同的值，则二阶接触就出现这种情况。与此相反的是，接触的阶恰好是奇次的，例如，一个通常的一阶接触，就意味着对所有用较小的数表示的 h 值，不管正的或者负的，$f(a+h)-g(a+h)$ 这个差都具有相同的符号；因此两条曲线在接触点的一个邻域内不相交叉。最简单的例子是一条曲线和它的切线的接触。仅在接触至少是二阶的一些点上切线才能与曲线交叉；而当接触的阶是偶数时它就真正穿过曲线。例如，一个通常的拐点，那里 $f''(x)=0$ 但 $f'''(x)\neq 0$。在接触的阶是奇数的点上，切线不会和曲线交叉，

1) 两条曲线的接触的阶是一个真正的几何关系，它是不受轴的转动的影响的，这一事实很容易用转轴公式来证实(参看第四章第 385 页)。

作为一个例子，在一个二阶导数不为零的曲线通常的点上，例如，曲线 $y=x^2$ 在原点处就是这样.

我们从第四章的第 384 页里知道，对于由函数 $y=g(x)$ 在 $x=a$ 的一个邻域里给定的在点 $x=a$ 处的曲率圆，我们不仅有 $g(a)=f(a)$ 和 $g'(a)=f'(a)$，而且也有 $g''(a)=f''(a)$. 因此曲率圆同时也是在所讨论的曲线的一点处的密切圆；就是说，在那个点处和曲线有二阶接触的一个圆.就拐点的极限情形而论，或一般说来，在这个点曲率是零而曲率半径是无穷大时，曲率圆就退化成切线. 就正常情形说来，当可讨论点处的接触不是高于二阶时，则曲率圆不仅仅与曲线接触，而且与它交叉(参看第382页图4.23).

最后要提一提，接触的阶恰恰是 m 可以这样描述：两条曲线有 $m+1$ 个无限接近的公共点；自然这样陈述的精确意义显然要涉及到极限过程. 事实上如果这些曲线共有 $m+1$ 个共同点 P, $P_1,\cdots,P_m$，令所有的点 P_i 趋向于 P，必要时可以仅更改曲线中的一个，那末极限状态必然如所期望的，就是两条曲线具有 m 阶接触. 例如，如果我们在曲线 C 上画一个圆通过三个点 P, P_1, P_2，而后令 P_1 和 P_2 趋向于 P，就能看到这圆趋向于 C 上 P 点处的曲率圆(参看第 461 页上的问题 4).

b. 关于相对极大值和相对极小值的理论

我们在第三章第 258 页已经看到，一个在 $x=a$ 点一阶导数为零的函数 $f(x)$，如果 $f''(a)$ 是负的，则在这个点有一个相对的极大值；如果 $f''(a)$ 是正的，则有一个相对的极小值. 因此，这些条件是发生极大值或极小值的充分条件. 它们决不是必要的；因为当 $f''(a)=0$ 时有三种可能：在所讨论的点上函数可以有极大值或极小值或二者都没有. 在点 $x=0$ 处函数 $y=-x^4$, $y=x^4$, 和 $y=x^3$ 给出了这三种可能的例子. 泰勒定理立即使我们能够对极大值或极小值的充分条件作一个一般的陈述. 我们只须展开函数 $f(a+h)$ 成 h 的幂，这时，问题的关键是看第一个非零项是 h 的偶次幂还是奇次幂. 在第一种情形下我们有一个极大值或一个

极小值，随 h 的系数是正的或者是负的而定；在第二种情形下我们有一个水平的扭转切线. 因而，既没有极大值也没有极小值. 用余项的公式读者可自行完成论证[1].

附　录　I

A.I.1 不能展成泰勒级数的函数的例

通过一个具有 $n+1$ 阶余项的泰勒级数来表示一个函数的可能性本质上依赖于这函数在所讨论的点处的连续性和可微性. 鉴于这个道理，$\log x$ 不能用 x 的幂表示成一个泰勒级数，而函数 $x^{\frac{1}{3}}$ 也一样，因为它的导数在 $x=0$ 点是无穷的.

若要一个函数可能展成一个无穷的泰勒级数，它的所有阶导数都必须在所讨论的点处存在；然而，这个条件决不是充分的. 一个函数的所有导数在整个区间存在并且连续，仍旧不一定能展成一个泰勒级数；也就是说，不管我们所要展开函数的区间如何小，泰勒定理的余项 R_n 未必随 n 增加而趋于零.

这个现象的一个重要而又简单的例子是函数

$$y=f(x)=e^{-1/x^2}\quad 对\quad x\neq 0,\quad f(0)=0.$$

这在第三章第 272 页的附录中我们已经考虑过. 这个函数和它的所有导数在每个区间甚至在 $x=0$ 点都是连续的，并且显而易见，在这个点所有导数都为零，即对 n 的每一个值 $f^{(n)}(0)=0$.（几何上就是说，直线 $y=0$ 在 $x=0$ 点和函数的曲线有无穷阶接触.）因此在泰勒展开式

$$f(0)+\frac{x}{1!}f'(0)+\frac{x^2}{2!}f''(0)+\cdots$$

里，不论 n 取什么值，近似多项式 $P_n(x)$ 的全部系数都是零. 因此

1) 然而，以前给出的必要和充分条件（第257页）在应用上更一般，更便利：只要一阶导数 $f'(x)$ 仅在有限个点上为零，在这些点中的一个发生极大值或极小值的充分和必要条件是，当曲线通过此点时，一阶导数 $f'(x)$ 变号.

余项始终等同于函数本身；而因为函数对除 $x=0$ 外的每一个其他 x 值是正的，所以当 n 增加时余项 R_n 不可能趋于零.

顺便说明一下，这个函数对于构造这样的函数是有用途的，它们显示了在直观上人们所意想不到的现象. 例如

$$g(x)=e^{-1/x^2}\sin(1/x),$$

补充上 $g(0)=0$，是一个在 $x=0$ 点所有阶导数都为零的函数；在 $x=0$ 邻近，$y=g(x)$ 的图象与 x 轴交叉无穷多次，并且振动无穷多次.

A.I.2 函数的零点和无限点

a. n 阶零点

一个函数 $f(x)$ 的泰勒展开式使我们能够刻划一个函数在一点 $x=a$ 处消失的阶的特征. 若 $f(a)=0$，$f'(a)=0$，$f''(a)=0$，$\cdots$，$f^{(n-1)}(a)=0$，而 $f^{(n)}(a)\neq 0$，我们就说这个函数 $f(x)$ 在 $x=a$ 恰有 n 重的零点，或者说它在 $x=a$ 为零的阶恰好是 n. 我们明确地假设，在这个点的邻域里函数至少 n 阶连续可微. 按照我们的定义，就意味着函数的泰勒级数在这点的一个邻域里能写成如下的形式

$$f(a+h)=\frac{h^n}{n!}F(h)=\frac{h^n}{n!}f^{(n)}(a+\theta h),\ 0<\theta<1,\quad(28)$$

当 $h\to 0$ 时因子 $F(h)=n!f(a+h)/h^n$ 趋向于一个不为零的极限 $f^{(n)}(a)$. 因此 $f(a+h)$ 当 $h\to 0$ 时与 h^n 具有同样的阶数，或如第三章第 268 页定义的意义下，以阶数 n 趋于零.

同样地，由具有拉格朗日形式的余项的泰勒定理展开导数 $f'(x)$，$f''(x)$，$\cdots$，$f^{(\nu)}(x)$，我们就得到一系列表达式

$$f'(a+h)=\frac{h^{n-1}}{(n-1)!}F_1(h)=\frac{h^{n-1}}{(n-1)!}f^{(n)}(a+\theta h),$$

$$\cdots\cdots\quad(29)$$

$$f^{(\nu)}(a+h)=\frac{h^{n-\nu}}{(n-\nu)!}F_\nu(h)=\frac{h^{n-\nu}}{(n-\nu)!}f^{(n)}(a+\theta h),$$

其中所有因子 θ 可以是不同的，而因子 F_1，F_2，$\cdots$，F_ν 当 $h\to 0$ 时连续地趋向于 $f^{(n)}(a)$. 因此 f' 变为零的阶数为 $n-1$，f'' 为

$n-2$ 阶，等等.

当然，在这些公式里，假定了 $f(x)$ 变为零的阶数是 $n \geqslant \nu$.

b. ν 阶无限

如果一个函数 $\phi(x)$，除了 $x=a$ 本身以外，在其邻域里所有的点上有定义，并且若 $\phi(x)=f(x)/g(x)$，其中在点 $x=a$ 分子不为零，但是分母具有一个 ν 重零点，我们就说函数 $\phi(x)$ 在 $x=a$ 点变成无穷的阶数为 ν. 如果在点 $x=a$ 分子有一个 μ 重零点且若 $\mu>\nu$，函数就在那里有一个 $(\mu-\nu)$ 重零点；如果 $\mu<\nu$，函数就在这个点上有一个 $(\nu-\mu)$ 重无限点.

这些定义与已经叙述过的(参看第 3.7 节)关于函数的特性的约定相符合.

A.I.3 不 定 式

我们现在更精确地来讨论形如 $\phi(x)=f(x)/g(x)$ 的“不定式”，其中 $f(x)$ 和 $g(x)$ 在同一点 $x=a$ 都变为零，例如函数 $(\sin x)/x$ 在 $x=0$ 时的情况. 我们将始终指定这样的函数在 $x=a$ 的值为:

$$\phi(a)=\lim_{h\to 0}\phi(a+h), \tag{30}$$

只要这个极限存在.

我们假设，出现的 f 和 g 的所有阶导数在包含点 $x=a$ 的一个区间里都是连续的. 此外我们假设，在点 $x=a$ 分母 $g(x)$ 变为零的阶数 ν 不高于分子 $f(x)$ 变为零的阶数，从而函数 $\phi(x)$ 在点 $x=a$ 处不致于变为无穷. 这时极限值就能用一个简单的法则来确定，这就是通常所说的洛必达法则. 这个法则可以这样叙述:

$$\phi(a)=\frac{f^{(\nu)}(a)}{g^{(\nu)}(a)}. \tag{31}$$

按照连续性的定义，函数 $\phi(x)$ 就在 $x=a$ 是连续的，并且在其他 $g(x)\neq 0$ 的地方也是连续的，所以 ϕ 就在关于 a 附近的

一个区间里都是连续的.

由 A.2 节的结果立即得到以下的证明：对 f 和 g 应用等式(28),我们看到函数 ϕ 在 a 的一个邻域里由关系式

$$\phi(a+h)=\frac{f(a+h)}{g(a+h)}=\frac{f^{(\nu)}(a+\theta h)}{g^{(\nu)}(a+\theta h)}$$

给出，于是由于分子和分母都是连续的，并且 $g^{(\nu)}(a)$ 不是零，就得到公式(31). 我们可以用以下的方法来表示最后这几个等式的意义：如果函数 $\phi(x)=f(x)/g(x)$ 的分子和分母在点 $x=a$ 一起变为零，就可以对分子和分母求导相同的次数，直到导数至少有一个在这点不为零，我们就能确定当 $x\to a$ 时的极限值. 如果我们在分母里遇到一个出现在分子前面不为零的导数，则分式趋于零. 如果我们在分子里遇到一个出现在分母前面不为零的导数，则分式的绝对值趋于无穷大.

我们于是得到了所谓"不定型" 0/0 的求值方法，即，用来确定当分子和分母趋于零时商的极限值的方法.

我们还可以用稍微不同的方法得到上述的结果，即不以泰勒定理为依据，而以广义中值定理（参看第 236 页）的证明为基础. 因此，假定在 a 的一个邻域里 $g'(x)\neq 0$，我们有

$$\frac{f(a+h)-f(a)}{g(a+h)-g(a)}=\frac{f'(a+\theta h)}{g'(a+\theta h)},$$

其中分子和分母中的 θ 是相同的. 特别是，当 $f(a)=g(a)=0$ 时，

$$\frac{f(a+h)}{g(a+h)}=\frac{f'(a+\theta h)}{g'(a+\theta h)}.$$

其中 θ 是在区间 $0<\theta<1$ 里的一个值，因而令 $k=\theta h$，就得到

$$\lim_{h\to 0}\frac{f(a+h)}{g(a+h)}=\lim_{k\to 0}\frac{f'(a+k)}{g'(a+k)},$$

只要后一个极限存在.

如果 $f'(a)=0=g'(a)$，用前面同样的方法进行，直到第一

次遇到这样的 μ，不再同时成立 $f^{(\mu)}(a)=0=g^{(\mu)}(a)$. 于是

$$\lim_{h\to 0}\frac{f(a+h)}{g(a+h)}=\lim_{l\to 0}\frac{f^{(\mu)}(a+l)}{g^{(\mu)}(a+l)}=\frac{f^{(\mu)}(a)}{g^{(\mu)}(a)},$$

两边是无穷大时的情形也包括在这个式子里.

例. 下列各个意义很显明的例子，都是洛必达法则的应用.

$$\lim_{x\to 0}\frac{\sin x}{x}=\frac{\cos 0}{1}=1;$$

$$\lim_{x\to 0}\frac{1-\cos x}{x}=\frac{\sin 0}{1}=0;$$

$$\lim_{x\to 0}\frac{e^{2x}-1}{\log(1+x)}=\lim_{x\to 0}\frac{2e^{2x}}{1/(1+x)}=2;$$

$$\lim_{x\to 0}\frac{1-\cos x}{x^2}=\lim_{x\to 0}\frac{\sin x}{2x}=\lim_{x\to 0}\frac{\cos x}{2}=\frac{1}{2}.$$

其它不定式. 我们进一步注意到其它所谓的不定式也能化成我们所考虑过的情形. 例如，

$$\frac{1}{\sin x}-\frac{1}{x}$$

当 $x\to 0$ 时的极限，就是趋于无穷的两个表达式的差的极限，或说是"不定式" $\infty-\infty$. 由于变换

$$\frac{1}{\sin x}-\frac{1}{x}=\frac{x-\sin x}{x\sin x},$$

我们立即得出一个表达式，当 $x\to 0$ 时，它的极限可以由我们的法则确定：

$$\lim_{x\to 0}\frac{1-\cos x}{x\cos x+\sin x}=\lim_{x\to 0}\frac{\sin x}{2\cos x-x\sin x}=0.$$

不定式的导数

如果 f 和 g 有足够高阶的连续导数，按照上述法则在 $x=a$ 处给表达式 $\phi(x)=f(x)/g(x)$ 以定义，则它不仅连续，而且连续可微.

要建立这一事实，我们只需要考虑 g 在 a 点有一阶零点的情

况就够了，即只要考虑 $g(a)=0$，$g'(a)\neq 0$ 的情况．对 $x\neq a$

$$\phi'(x)=\frac{g(x)f'(x)-f(x)g'(x)}{(g(x))^2}=\frac{Z(x)}{N(x)},$$

在这里分子和分母又一次在 $x=a$ 同时为零，因此应用我们的法则就能得到极限值：

$$\lim_{x\to a}\phi'(x)=\lim_{x\to a}\frac{Z'(x)}{N'(x)}.$$

显然

$$d(N(x))/dx=2g(x)g'(x),$$
$$d(Z(x))/dx=g(x)f''(x)-f(x)g''(x),$$

二者在 $x=a$ 又一次为零．因

$$\lim_{x\to a}\phi'(x)=\lim_{x\to a}\frac{Z''(x)}{N''(x)},$$

再一次应用洛必达法则，并且注意到 $N''(x)=2g(x)g''(x)+2(g'(x))^2$ 在 $x=a$ 不为零，我们就得到

$$\lim_{x\to a}\phi'(x)=\frac{g'(a)f''(a)-f'(a)g''(a)}{2(g'(a))^2},$$

而这个极限确实是 $\phi'(x)$在$x=a$的导数（参看第三章第 216 页）．

同样，这个法则对 $x\to\infty$ 时的不定式也成立．譬如令 $f(x)$ 和 $g(x)$ 是两个函数，对于这两个函数 $\lim\limits_{x\to\infty}f(x)=\lim\limits_{x\to\infty}g(x)=0$，而 $\lim\limits_{x\to\infty}f'(x)$ 和 $\lim\limits_{x\to\infty}g'(x)$ 存在且 $\neq 0$．则

$$\lim_{x\to\infty}\frac{f(x)}{g(x)}=\frac{\lim\limits_{x\to\infty}f'(x)}{\lim\limits_{x\to\infty}g'(x)}.$$

证明仍可由微商中值定理推出．

*A.I.4 各阶导数都不为负的函数的泰勒级数的收敛性

我们补上一个一般的定理，研究关于所有导数都不为负的函数 $f(x)$ 的泰勒展开的收敛性．

考虑一类函数 $f(x)$，在闭区间 $a\leqslant x\leqslant b$ 上任意阶可导，并

在这区间上它的各阶导数不为负，即

$$f^{(\nu)}(x) \geqslant 0, \quad \nu = 1, 2, \cdots.$$

我们将证明：对每一个这样的函数，对应的 $f(x+h)$ 关于 h 的幂的泰勒展开式收敛，并且当 x 和 $\xi = x + h$ 位于开区间(a, b)内，而 $|h| < b - x$ 时，级数就表示 $f(x+h)$ 的值.

为了证明，我们首先注意到假设 $f'(x) \geqslant 0$，因而

$$0 \leqslant f(x) - f(a) = \int_a^x f'(\xi)d\xi$$

$$\leqslant \int_a^b f'(\xi)d\xi = f(b) - f(a) = M.$$

其次，当 x 和 $\xi = x + h$ 位于 a 到 b 的区间内时，我们可以写成

$$f(x+h) - f(x) = hf'(x) + \cdots + \frac{f^{(n)}(x)}{n!}h^n + R_n.$$

首先假设 $h > 0$，或者 $x < \xi < b$. 这样右边所有的项就都不是负的了[1]，所以每一项都不大于左边的值，或不大于 M；于是

$$0 \leqslant \frac{f^{(n)}(x)}{n!} \leqslant \frac{M}{h^n} = \frac{M}{(\xi - x)^n}.$$

当 $\xi \to b$ 时由此得出

$$\frac{f^{(n)}(x)}{n!} \leqslant \frac{M}{(b-x)^n}. \tag{32}$$

现在，对余项应用柯西公式(第 472 页(19)) 我们知道在区间 $0 < \theta < 1$ 内存在某个 θ，使得

$$0 \leqslant R_n = \frac{(1-\theta)^n}{n!} h^{n+1} f^{(n+1)}(n + \theta h)$$

$$\leqslant \frac{h^{n+1}(n+1)(1-\theta)^n M}{(b - x - \theta h)^{n+1}}.$$

因为 $\xi = x + h < b$，我们可以选择一个正数 p 使得

$$0 \leqslant h \leqslant \frac{b-x}{1+p} \quad 或 \quad b - x - \theta h \geqslant h(1 + p - \theta).$$

1) 对于 R_n 而言，这可从柯西或拉格朗日公式和假设 $f^{(n+1)} \geqslant 0$ 立即得出.

于是我们有

$$0 \leqslant R_n \leqslant \frac{Mh^{(n+1)}(n+1)(1-\theta)^n}{h^{n+1}(1+p-\theta)^{n+1}}$$

或

$$0 \leqslant R_n \leqslant \frac{M(n+1)}{(1+p-\theta)}\left(\frac{1-\theta}{1-\theta+p}\right)^n$$

$$\leqslant \frac{M(n+1)}{p}\frac{1}{(1+p)^n},$$

因为

$$\frac{1-\theta}{1-\theta+p} = \frac{1}{1+p/(1-\theta)} \leqslant \frac{1}{1+p} < 1.$$

我们知道（第一章第 73 页）当 n 增大时 $(n+1)/(1+p)^n$ 趋于零，所以当 $0 \leqslant h < b-x$ 时，R_n 随 n 的增加而趋于零，于是对于 $h>0$ 泰勒级数趋于这个函数 f.

对于负的 h，R_n 随着 n 增加而趋于零这个事实，可用 R_n 的拉格朗日形式（第 473 页(21)）推出：

$$|R_n| = \frac{1}{(n+1)!}|h^{n+1}||f^{(n+1)}(x-\theta|h|)|.$$

现在 $f^{(n+2)}$ 是非负的，因此 $f^{(n+1)}$ 是单调减小的. 于是应用上述的估计式 (32) 就得出

$$\frac{f^{(n+1)}(x-\theta|h|)}{(n+1)!} \leqslant \frac{f^{(n+1)}(x)}{(n+1)!} \leqslant \frac{M}{(b-x)^{n+1}}.$$

所以 $$|R_n| \leqslant \left(\frac{|h|}{b-x}\right)^{n+1}M,$$

从而当 $0<-h<b-x$ 时，R_n 随 n 增大而趋于零.

于是对于 $a \leqslant x < b$ 的任意点 x，一旦 $|h|<b-x$ 和 $h>-(x-a)$，则 $f(a+h)$ 的依 h 的幂而展开的泰勒级数中的 R_n 便趋于零.

我们指出，如果我们假设不等式 $f^{(\nu)}(x) \geqslant 0$ 仅对所有足够大的 ν，譬如 $\nu > N$（N 是某整数）成立，而当 $\nu \leqslant N$ 时 $f^{(\nu)}(x)$ 的符号可以是任意的，那末我们的结果仍是正确的. 为了证明这个

事实，我们仅需用函数

$$g(x) = f(x) + M(x - a + 1)^N$$

来代替证明中的 f，其中M是某个正的常数．这样对于 $\nu > N$ 有 $g^{(\nu)}(x) = f^{(\nu)}(x) \geqslant 0$． 而对 $\nu \leqslant N$ 有 $g^{(\nu)}(x) = f^{(\nu)}(x) + MN(N-1)\cdots(N-\nu+1)(x-a+1)^{N-\nu} \geqslant f^{(\nu)}(x) + M$．如果 M 选择得足够大则对所有的 ν 都有 $g^{(\nu)}(x) \geqslant 0$． 这就证明了 $g(x)$ 能被展成 x 的幂，而对于和 g 仅差一个多项式的函数 f 当然也有同样的结果．

二项式级数的定理(第480页)是这个结果的一个直接的推论：我们稍微改变一下符号，代替 $(1+x)^\alpha$ 我们首先考虑函数 $\phi(x) = (1-x)^\alpha$，那么ϕ的各阶导数由

$$\phi^{(\nu)}(x) = (-1)^\nu \binom{\alpha}{\nu} (1-x)^{\alpha-\nu} \nu!$$

给出．因为这个二项系数

$$\binom{\alpha}{\nu} = \frac{\alpha(\alpha-1)\cdots(\alpha-\nu+1)}{\nu!}$$

一旦到了 $\alpha - \nu$ 是负的时候就开始交替变号，这时我们就看到，或者函数 $\phi(x)$或者 $-\phi(x)$属于这类函数，当我们限定取 $x < 1$ 的 x 值时，从某一阶开始就具有非负的导数．这样，对于 $a = -1$，$b = 1$，$x = 0$，和 $|h| < b - x = 1$，我们的一般的定理证明了

$$(1-h)^\alpha = \sum_{\nu=0}^{\infty} (-1)^\nu \binom{\alpha}{\nu} h^\nu.$$

如果在这里我们把 $-h$ 写成 x，我们就得到对任意指数α 和绝对值小于 1 $(-1 < x < 1)$ 的任意 x 的二项展开式

$$(1+x)^\alpha = \sum_{\nu=0}^{\infty} \binom{\alpha}{\nu} x^\nu = 1 + \alpha x + \frac{\alpha(\alpha-1)}{1\cdot 2} x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{1\cdot 2\cdot 3} x^3 + \cdots.$$

附录II 插值法

*A.II.1 插值问题．唯一性

泰勒多项式 $P_n(x)$ 是以这样一种方式逼近于函数 $f(x)$ 的：即 $f(x)$ 和 $P_n(x)$ 的图象在点 a 有 n 阶接触，或者是这样的方式，$f(x)$ 和 $P_n(x)$ 在“无限接近”点 a 处有 $n+1$ 点重合．我们可以把这个具有横坐标 a 的点“分解”成为具有横坐标 $x_0, x_1, \cdots, x_n$ 的不同的 $n+1$ 个点，并且寻求一个在这些点上与 $f(x)$ 重合的近似于 $f(x)$ 的 n 次多项式 $\phi(x)$．这个多项式原来是按照一个线性方程组唯一地确定的．对于所有的 i 取极限 $x_i \to x$，我们就回到了这个泰勒多项式．但是，“插值法”（即这个按照在不同的点上与 $f(x)$ 相重合所得到的近似多项式）在许多应用上是特别重要的．以下的讨论将给出插值理论的一个简短的说明．

我们考虑以下的问题：决定一个 n 次多项式 $\phi(x)$，使得它在给定的 $n+1$ 个不同的点 $x_0, x_1, \cdots, x_n$ 上取得 $n+1$ 个给定的值 $f_0, f_1, \cdots, f_n$，即

$$\phi(x_0) = f_0, \quad \phi(x_1) = f_1, \cdots, \phi(x_n) = f_n.$$

如果数 f_i 是函数（可能不是初等的）$f(x)$ 在点 x_i 所取的值 $f_i = f(x_i)$，那么多项式 $\phi(x)$ 就称为函数 $f(x)$ 在 $x_0, x_1, \cdots, x_n$ 点的 n 次**插值多项式**．

这样的 n 次多项式至多只能有一个，因为如果有两个不同的这样的多项式 $\phi(x)$ 和 $\psi(x)$，那么它们的差 $D(x)=\phi(x)-\psi(x)$ 将是一个 m 次多项式，其中 $0 \leqslant m \leqslant n$，它有 $n+1$ 个不同的根，根据初等代数这是不可能的[1]．

---

1) 因为否则我们会有

$$D(x) = C_0(x - x_1)(x - x_2)\cdots(x - x_m), \quad c_0 \neq 0,$$

因为 $x_1, \cdots, x_m$ 是 $D(x)$ 的零点；但另一方面，由于 $D(x_0) = 0$，

$$C_0(x_0 - x_1)(x_0 - x_2)\cdots(x_0 - x_m) = 0,$$

与 $x_0, x_1, \cdots, x_m$ 都是不同的根相矛盾．

我们也能用另一方法证明插值多项式的唯一性，这个方法建立在广义罗尔定理的基础上.

广义罗尔定理. 如果一个函数 $F(x)$ 在一个区间里直到 n 阶连续可微，并且在这个区间上至少在 $n+1$ 个不同的点 $x_0, x_1, \cdots, x_n$ 上变为零，那么在这个区间内部一定存在一点 ξ 使得 $F^{(n)}(\xi)=0$.

证明：这个一般的定理很容易从 $n=1$ 的特殊情形得到，这个特殊情形就是在第 186 页上已经证明过的罗尔定理. 设这些数 $x_0, x_1, \cdots, x_n$ 是按逐个增大的顺序排列的，那么由中值定理（或罗尔定理），一阶导数 $F'(x)$ 在这 n 个子区间 (x_i, x_{i+1}) 的每一个内部至少一次为零. 应用同样的考虑到 $F'(x)$，$F'(x)$ 的零点组成的区间告诉我们 $F''(x)$ 在 $n-1$ 个点上为零；反复运用这个论证，断言就被证明了.

现在回到前面考虑的问题，我们对差

$$F(x)=D(x)=\phi(x)-\psi(x)=d_0x^n+d_1x^{n-1}+\cdots+d_n$$

应用这个定理，假设这个差在 $n+1$ 个点上变为零. 我们得到一点 ξ，在这个点差的 n 阶导数为零：$D^{(n)}(\xi)=0$. 然而这就是 $n!d_0$，所以 $d_0=0$ 而差是一个至多 $n-1$ 次的多项式，在 $n+1$ 个点上为零. 再一次应用罗尔定理，就得到 $d_1=0$，等等，正如我们所断言的 $D(x)$ 恒为 0.

这些考虑能扩充到 x_i 不是全部相异的情形，也许 r 个 x_i 的值相同；即，$x_0=x_1=\cdots=x_{r-1}$. 这样对 $x=x_0$ 来说，在插值问题里我们将要求 $\phi(x)$ 以及导数 $\phi'(x), \cdots, \phi^{(r-1)}(x)$ 在 $x=x_0$ 取预先给定的值，而对于其他的点 x_ν 也取相应的值. 于是这个多项式 $D(x)$ 的形式是 $c(x-x_0)^r(x-x_r)\cdots$. 广义罗尔定理和唯一性定理以及它们的证明在这种情形下一律保持不变.

A.II.2 解的构造. 牛顿插值公式

现在我们将构造一个 n 次插值多项式 $\phi(x)$，使得 $\phi(x_0)=f_0, \cdots, \phi(x_n)=f_n$. 为了分层次去构造它，我们将从一个常数 f_0

开始，这是一个零次多项式 $\phi_0(x)$，它对所有 x（其中也对 $x=x_0$）取值 $A_0=f_0$. 对于它我们加上一个在 $x=x_0$ 时变为零的一次多项式 $A_1(x-x_0)$；由此我们确定 A_1，使得和数当 $x=x_1$ 时恰取值 f_1. 所得到的一次多项式称为 $\phi_1(x)$. 现在我们对 $\phi_1(x)$ 加上一个当 $x=x_0$ 和 $x=x_1$ 时变为零的二次多项式，因而其形式为 $A_2(x-x_0)(x-x_1)$，这样的加法不会改变在这两个点的性质；其中因子 A_2 这样来确定，使得所得到的二次多项式 $\phi_2(x)$ 在 $x=x_2$ 取给定的数值，这就是 f_2. 把这个手续继续下去，直到所有的点都被达到为止，我们就得到多项式

$$\phi(x)=\phi_n(x)=A_0+A_1(x-x_0)+A_2(x-x_0)(x-x_1)+\cdots+A_n(x-x_0)\cdots(x-x_{n-1}). \tag{33}$$

我们得出这个表达式 $\phi(x)$ 里的系数 A_i 的方法只要依次代入 $x=x_0$, $x=x_1$, $\cdots$, $x=x_n$ 就可以清楚地看出来，于是就得到这样一组 $n+1$ 个方程：

$$\begin{aligned}
f_0&=A_0\\
f_1&=A_0+A_1(x-x_0)\\
f_2&=A_0+A_1(x_2-x_0)+A_2(x_2-x_0)(x_2-x_1)\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
f_n&=A_0+A_1(x_n-x_0)+\cdots+A_n(x_n-x_0)\\
&\quad\times(x_n-x_1)\cdots(x_n-x_{n-1}).
\end{aligned} \tag{34}$$

很清楚，我们可以相继地确定出这些系数 A_0, A_1, $\cdots$, A_n，使之满足这些方程. 因此插值多项式可以用这样的方法构造出来.

当值 x_ν 是等距 $x_\nu=x_{\nu-1}+h$ 时，结果就能明白写成更漂亮的方式. 关于 A_i 的方程现在变成

$$\begin{aligned}
f_0&=A_0\\
f_1&=A_0+hA_1\\
f_2&=A_0+2hA_1+2!h^2A_2\\
f_3&=A_0+3hA_1+3\cdot 2h^2A_2+3!h^3A_3\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
f_n&=A_0+nhA_1+\cdots+\frac{n!}{(n-i)!}h^iA_i+\cdots+n!h^nA_n.
\end{aligned} \tag{35}$$

这些解可以很容易地表示成 f 的逐次差分：给定任意项的序列(有穷的或无穷的) $f_0, f_1, f_2, \cdots$ 我们称这些表达式

$$\Delta f_0 = f_1 - f_0, \quad \Delta f_1 = f_2 - f_1, \quad \Delta f_2 = f_3 - f_2, \cdots$$

为 f_k 的一次差分．对 Δf_k 的序列再一次应用差分的手续,我们就得到表达式

$$\Delta^2 f_0 = \Delta f_1 - \Delta f_0, \quad \Delta^2 f_1 = \Delta f_2 - \Delta f_1, \quad \Delta^2 f_2 = \Delta f_3 - \Delta f_2, \cdots,$$

即,

$$\Delta^2 f_0 = f_2 - 2f_1 + f_0, \quad \Delta^2 f_1 = f_3 - 2f_2 + f_1, \cdots,$$

称为 f_k 的二次差分．n 次差分 $\Delta^n f_k$ 递归地定义为 $\Delta^{n-1} f_{k+1} - \Delta^{n-1} f_k$；若要直接用 f_k 表示,便由公式

$$\Delta^n f_k = f_{k+n} - \binom{n}{1} f_{k+n-1} + \binom{n}{2} f_{k+n-2} - + \cdots + (-1)^n f_k \tag{36}$$

给出．此式的一个简单的归纳证明留给读者．借助这一术语，系数 A_ν 可以写成如下的形式[1]

$$A_\nu = \frac{1}{\nu!} h^{-\nu} \Delta^\nu f, \tag{37}$$

这可以用归纳法验证．

牛顿插值公式．令 $\xi = (x - x_0)/h$，我们有 $(x - x_\nu) = h(\xi - \nu)$，这时表达式 $(x - x_0)(x - x_1)\cdots(x - x_\nu)$ 取这样的形式：$\xi(\xi - 1)\cdots(\xi - \nu)h^{\nu+1}$．由 (33) 和 (37) 我们就得到多项式 $\phi(x)$ 的牛顿插值公式[2]：

---

1) 我们还必须验证,由 (37) 给定的值 A_ν 满足方程 (35)；即对任意的序列 $f_0, f_1, f_2, \cdots$，等式

$$f_k = f_0 + \binom{k}{1}\Delta f_0 + \binom{k}{2}\Delta^2 f_0 + \cdots + \binom{k}{k}\Delta^k f_0$$

恒被满足．假设这对于一个确定的 k 是正确的,我们必须证明

$$\begin{aligned} f_{k+1} &= f_1 + \binom{k}{1}\Delta f_1 + \binom{k}{2}\Delta^2 f_1 + \cdots \\ &= (f_0 + \Delta f_0) + \binom{k}{1}(\Delta f_0 + \Delta^2 f_0) + \binom{k}{2}(\Delta^2 f_0 + \Delta^3 f_0) + \cdots \\ &= f_0 + \binom{k+1}{1}\Delta f_0 + \binom{k+1}{2}\Delta^2 f_0 + \cdots \end{aligned}$$

这就是恒等式对于 $k+1$ 的情形．

2) 如同第 481 页,这里我们对一般的 ξ 和正整数 k 把二项系数 $\binom{\xi}{k}$ 定义为

$$\binom{\xi}{k} = \xi(\xi - 1)\cdots(\xi - k + 1)/k!$$

$$\phi(x)=\phi(x_0+\xi h)=f_0+\binom{\xi}{1}\Delta f_0+\binom{\xi}{2}\Delta^2 f_0+\cdots+\binom{\xi}{n}\Delta^n f_0.$$

如果 $f_0, f_1, f_2, \cdots$ 是一个函数 $f(x)$ 在点 $x_0, x_1, x_2, \cdots$ 的值，其中 f 直到 n 阶可连续求导，则 $\Delta^\nu f_0/h^\nu$ 是导数 $f^{(\nu)}(x_0)$ 的一个近似；我们将在第 501 页上证明

$$\lim_{h\to 0}\frac{1}{h^\nu}\Delta^\nu f_0=f^{(\nu)}(x_0).$$

又因为

$$\lim_{h\to 0} h^k\binom{\xi}{k}=\frac{(x-x_0)^k}{k!},$$

我们看到，在这种情形下，当 $h\to 0$ 时 $\phi(x)$ 趋向于泰勒多项式 $P_n(x)$.

我们指出，对于给定的函数 f，如果前 r 个值 $x_0, x_1, \cdots, x_{r-1}$ 相重合，并且对应值 $f_0, f_0', \cdots, f_0^{(r-1)}$ 是预先指定给 $\phi(x_0), \phi'(x_0), \cdots, \phi^{(r-1)}(x_0)$ 的，它与 $f(x_0)$，$f'(x_0)$，$\cdots, f^{(r-1)}(x_0)$ 的值相一致，那末用同样的方式构造这个插值多项式也是可能的. 对于 $\phi(x)$ 我们写成

$$\begin{aligned}\phi(x)&=A_0+A_1(x-x_0)+A_2(x-x_0)^2\\&\quad+\cdots+A_r(x-x_0)^r+A_{r+1}(x-x_0)^r(x-x_r)+\cdots;\end{aligned}$$

然后依据下列方程我们可以顺序确定 A_ν:

$$f_0=A_0\quad f_0'=A_1\quad f_0''=2A_2$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$f_0^{(r-1)}=(r-1)!A_{r-1}$$

$$f_r=A_0+A_1(x_r-x_0)+\cdots+A_r(x_r-x_0)^r$$

$$\begin{aligned}f_{r+1}&=A_0+A_1(x_{r+1}-x_0)+\cdots\\&\quad+A_r(x_{r+1}-x_0)^r+A_{r+1}(x_{r+1}-x_0)^r(x_{r+1}-x_r)\end{aligned}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots.$$

A.II.3 余项的估计

对于以上考虑的插值多项式，根本不涉及 $f_0, f_1, \cdots, f_n$ 这些值原来是怎样给定的. 例如，若这些值是从物理观测中得到的，构造插值多项式的问题仍然能完全解决，通过 $\phi(x)$ 给于我们一个简单

而光滑的函数，它对所有的 x 有定义，并在给定点上取观测到的值，它可以用来"预言" $f(x)$ 在其他点上的近似值. 然而，如果在给定的 $n+1$ 个点 x_k 上取给定的值 f_k 的函数 $f(x)$ 对中间值 x 也被定义了，那么我们就面临着一个估计插值的误差 $R(x)=f(x)-\phi(x)$ 的新问题. 最初我们只知道 $R(x_0)=R(x_1)=R(x_2)=\cdots=R(x_n)=0$. 为了能够说得更多些，我们必须给影响余项 $R(x)$ 的函数 $f(x)$ 的性状以更多的假设. 因此我们假定在所考虑的区间中 $f(x)$ 至少有 $n+1$ 阶连续导数.

首先我们注意到，对任意选择的常数 c，函数

$$K(x)=R(x)-c(x-x_0)(x-x_1)\cdots(x-x_n)$$

在这 $n+1$ 个点 $x_0,\cdots,x_n$ 上都为零. 现在选择一个不同于 $x_0,x_1,\cdots,x_n$ 的任意值 y. 这时，我们能够确定 c 使得 $K(y)=0$，即，

$$c=\frac{R(y)}{(y-x_0)(y-x_1)\cdots(y-x_n)}.$$

那么就有 $n+2$ 个点，在这些点上 $K(x)$ 为零. 对 $K(x)$ 我们用前面用过的广义罗尔定理，据此，我们知道，在 $x_0,x_1,\cdots,x_n,y$ 的最大值和最小值之间存在一个值 $x=\xi$，使得 $K^{(n+1)}(\xi)=0$. 因为 $R(x)=f(x)-\phi(x)$，而 ϕ 作为一个 n 次多项式，它的 $n+1$ 阶导数等于零，我们有

$$f^{(n+1)}(\xi)-c(n+1)!=0.$$

这里 $(n+1)!$ 是 $(x-x_0)\cdots(x-x_n)$ 的 $n+1$ 阶导数. 这样我们就得到 c 的第二个表达式 $c=f^{(n+1)}(\xi)/(n+1)!$，它包含 ξ，且以某种方式依赖于 y. 我们现在利用方程 $K(y)=0$，在这个方程里 y 完全是任意的，因而可用 x 替换，从而得到表达式

$$R(x)=\frac{(x-x_0)(x-x_1)\cdots(x-x_n)}{(n+1)!}f^{(n+1)}(\xi),\qquad(38)$$

其中 ξ 是位于 $x,x_0,x_1,\cdots,x_n$ 这些点的最大值和最小值之间的某个值.

这样，对于一个给定的函数 $f(x)$，一般的插值问题就完全地

解决了. 对于函数 $f(x)$, 我们有表达式

$$f(x) = A_0 + A_1(x - x_0) + A_2(x - x_0)(x - x_1) + \cdots + A_n(x - x_0)(x - x_1)\cdots(x - x_{n-1}) + R_n, \quad (39)$$

其中系数 $A_0, A_1, \cdots, A_n$ 能够从 f 在 $x_0, x_1, \cdots, x_n$ 的值按照第 497 页上的递归公式(34)依次求出来,而余项 R_n 的形式为

$$R_n = \frac{(x - x_0)(x - x_1)\cdots(x - x_n)}{(n + 1)!} f^{(n+1)}(\xi), \quad (40)$$

其中 ξ 是在 $x, x_0, x_1, \cdots, x_n$ 这些值的最大值和最小值之间的一个适当的数.

如果我们对 $f(x)$ 取相应的公式(39),把 n 换成 $n - 1$, 而后相减,我们就得到

$$A_n(x - x_0)(x - x_1)\cdots(x - x_{n-1}) + R_n - R_{n-1} = 0.$$

对于 $x = x_n$, 我们有 $R_n = 0$, 因此系数 A_n (应用(40), 把 n 换成 $n - 1$)表示为

$$A_n = \frac{f^{(n)}(\xi)}{n!},$$

其中 ξ 位于 $x_0, x_1, \cdots, x_n$ 的最大值和最小值之间. 对 A_{n-1}, $A_{n-2}, \cdots, A_0$ 存在着同样的表示方法. 这样我们认识到, 若点 $x_0, x_1, \cdots, x_n$ 同时趋向于同一个点 (也许就是原点), 那么我们的插值公式 (39) 就能一项一项地变成为具有拉格朗日形式余项 [第 473 页(21)]的泰勒公式 [第 475 页(27a)]. 这样泰勒公式可视为牛顿插值公式的极限情形.

这个公式使我们能够对通常用于几何中的一个表达式给出一个确切的含义. 在一点和一条给定的曲线相接触的 n 阶密切抛物线, 称为在这点和这条给定的曲线有" $n + 1$ 个依次相邻的共同点". 实际上,如果我们找到一条抛物线和该曲线有 $n + 1$ 个共同点,然后使这些点收缩到一起,我们就得到这条密切抛物线. 用解析语言说, 这正好相当于从内插多项式转变到泰勒多项式. 用这同样的方式,我们就能够刻划任意曲线的密切. 例如,曲率圆是一个与给定的曲线有三个相邻共同点的圆.

设一函数在某些确定点的值是已知的，由插值多项式可以期望给出这个函数的其他值，并使在这些确定点之间的值具有较高的精确度（这时不但 $|f^{(\nu+1)}(\xi)|$ 而且 $|x-x_i|$ 都是有界的）．若 x 值位于点 $x_0, x_1, \cdots, x_n$ 的所有区间以外，我们就称之为外插法，按照这样的一个外插法，只要点 x 充分地靠近给定的这些点，我们将会得到较好的符合．在某种意义上，泰勒公式相当于一个完全的外插法；一般说来，它仅在一点的邻域里适用．

A.II.4 拉格朗日插值公式

作为结束，我们用一个稍微不同的属于拉格朗日的公式来解决插值问题，这个公式不同于牛顿插值公式，因为这里每一个单项仅包含一个给定的函数值．并且这个公式很明显地给出了 $\phi(x)$，并不需要求解任何递归公式．为了简单起见，相应于给定的诸点 x_ν 我们引进 $n+1$ 次多项式

$$\psi(x)=(x-x_0)(x-x_1)\cdots(x-x_n),$$

按照乘法规则求导数，然后逐次以值 $x_0, x_1, \cdots, x_n$ 代 x，我们就得到关系式

$$\psi'(x_0)=(x_0-x_1)(x_0-x_2)\cdots(x_0-x_n),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\psi'(x_\nu)=(x_\nu-x_0)(x_\nu-x_{\nu-1})(x_\nu-x_{\nu+1})\cdots(x_\nu-x_n),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\psi'(x_n)=(x_n-x_0)(x_n-x_1)\cdots(x_n-x_{n-1}).$$

我们注意到

$$\frac{\psi(x)}{(x-x_\nu)\psi'(x_\nu)}=\frac{(x-x_0)\cdots(x-x_{\nu-1})(x-x_{\nu+1})\cdots(x-x_n)}{(x_\nu-x_0)\cdots(x_\nu-x_{\nu-1})(x_\nu-x_{\nu+1})\cdots(x_\nu-x_n)}$$

是一个 n 次多项式，在 $x=x_\nu$ 处其值为 1，而在其余的点 x_i 处其值为 0；于是立即可见，表达式

$$\phi(x)=\psi(x)\left[\frac{f_0}{(x-x_0)\psi'(x_0)}+\frac{f_1}{(x-x_1)\psi'(x_1)}+\cdots+\frac{f_n}{(x-x_n)\psi'(x_n)}\right]\quad(41)$$

就是所要求的插值多项式．这就是拉格朗日插值公式．

问　　题

5.4b 节，第 474 页

1．用数学归纳法给出第 475 页余式(27)的一个完全的正式的推导．

2．(泰勒定理的一个不同的证明)

(a) 如果 $g(h)$ 对 $0 \leqslant h \leqslant A$ 直到 $n+1$ 阶连续可微，并且

$$g(0) = g'(0) = \cdots g^{(n)}(0) = 0,$$

而在 $[0, A]$ 上 $|g^{(n+1)}(h)| \leqslant M$，$M$ 是常数，求证对这区间里所有的 h 都有 $|g^{(n)}(h)| \leqslant Mh$，$|g^{(n-1)}(h)| \leqslant Mh^2/2!$，$\cdots$，$|g^{(n-i)}(h)| \leqslant Mh^i/i!$，$\cdots$，$|g(h)| \leqslant Mh^n/n!$．

(b) 设 $f(x)$ 在 $a \leqslant x \leqslant b$ 上是一个充分可微的函数，$T_n(h)$ 是 $f(x)$ 在 $x = a$ 点的泰勒多项式．试对于函数 $g(h) = R_n = f(a+h) - T_n(h)$ 应用(a)的结果去得到对余项带有粗略估计的泰勒公式．

3．设 $f(x)$ 在区间 $a \leqslant x \leqslant b$ 连续可微，并设对每一个 x 的值 $f''(x) \geqslant 0$．如果 ξ 是区间内的任意点，那么，曲线无论什么地方都不会落在 $x = \xi$，$y = f(\xi)$ 点的切线的下面．

4．对 $f(x+h) - f(x) = \int_0^h f'(x+\tau)d\tau$ 应用分部积分公式来推演余式 R_n 的积分公式．

5．对公式

$$R_n = \frac{1}{n!}\int_0^h (h-\tau)^n f^{(n+1)}(x+\tau)d\tau$$

进行分部积分，从而得到

$$R_n = f(x+h) - f(x) - hf'(x) - \cdots - \frac{h^n}{n!}f^{(n)}(x).$$

*6．假设用某方法对函数 $f(x)$ 得到了级数，即

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + R_n(x),$$

其中 $a_0, a_1, \cdots, a_n$ 是常数，$R_n(x)$ 是 n 次连续可微的，并且当 $x \to 0$ 时，$R_n(x)/x^n \to 0$．证明 $a_k = (f^k(0)/k!)$ $(k = 0, \cdots, n)$，即，证明这个级数是一个泰勒级数．

5.5 节，第 477 页

1．对下列函数在 $x = 0$ 的邻域里求泰勒级数的不为零的前四项

(a) $x\cot x$,　　　　(d) $e^{\sin x}$,

(b) $\dfrac{\sqrt{\sin x}}{\sqrt{x}}$,　　　　(e) e^{e^x},

(c) $\sec x$,　　　　(f) $\log\sin x - \log x$.

2. 对 $\arcsin x$ 在 $x=0$ 的邻域里用

$$\arcsin x = \int_0^x \frac{dt}{\sqrt{1-t^2}}$$

求泰勒级数. 与 3.2 节习题 2 比较.

*3 对 $\sin^2 x$ 在 $x=0$ 的邻域里,用 $\sin x$ 的级数乘它自己来求泰勒级数的不为零的前三项. 证明这种求法是合理的.

*4. 对 $\tan x$ 在 $x=0$ 的邻域里,利用关系 $\tan x = \dfrac{\sin x}{\cos x}$,求泰勒级数的不为零的前三项. 证明这种求法是合理的.

*5. 对 $\sqrt{\cos x}$ 在 $x=0$ 的邻域里,对 $\cos x$ 应用二项式定理,证明这种求法是合理的.

*6. 求 $(\arcsin x)^2$ 的泰勒级数. 与 3.2 节问题 2 比较.

7. 对下列函数在 $x=0$ 的邻域里求泰勒级数:

(a) $\sinh^{-1} x$,　　(b) $\int_0^x e^{-t^2}dt$,　　(c) $\int_0^x \dfrac{\sin t}{t}\,dt$.

*8. 估计问题 7 里使用级数的前 n 项所产生的误差.

9. 在 3.14a 节里椭圆函数 $s(u)$ 定义为椭圆积分

$$u(s) = \int_0^s \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}$$

的反函数. 求 $s(u)$ 的泰勒展开式到 5 次项.

10. 计算下列极限:

(a) $\lim\limits_{x\to\infty} x\left[\left(1+\dfrac{1}{x}\right)^x - e\right]$,　　(d) $\lim\limits_{x\to 0}\left(\dfrac{\sin x}{x}\right)^{1/x^2}$,

(b) $\lim\limits_{x\to\infty}\left\{\dfrac{e}{2}x + x^2\left[\left(1+\dfrac{1}{x}\right)^x - e\right]\right\}$,　　(e) $\lim\limits_{x\to\infty}\left(\dfrac{\sin x}{x}\right)^{1/x^2}$.

*(c) $\lim\limits_{x\to\infty} x\left[\left(1+\dfrac{1}{x}\right)^x - e\log\left(1+\dfrac{1}{x}\right)^x\right]$,

*11. 求 $[1+(1/x)]^x$ 依 $\dfrac{1}{x}$ 的幂展开的泰勒级数的前三项.

*12 两颗异性的带电粒子 $+e$, $-e$, 当其间距离 d 很小时形成一个电偶

极子，其偶极矩 $M=cd$. 试证明

(a) 在位于偶极子轴上，而与偶极子的中心距离为 r 的一点处的电位能是 $(M/r^2)(1+\varepsilon)$，其中 ε 近似等于 $d^2/4r^2$.

(b) 在位于偶极子的中垂线上的一点处电位能是零.

(c) 在一个相对于偶极子的中心和轴的电位能是 $[M\cos(\theta/r^2)][1+\varepsilon]$，其中 ε 近似等于 $(d^2/8r^2)(5\cos^2\theta-3)$.

(单个的点电荷 q 在距这个电荷的距离为 r 的点处的电位能是 q/r; 几个电荷的电位能是单个电荷的电位能的和).

5.6 节，第 482 页

1. 若 $f(a)=0$ 而 $f(x)$ 在 $x=a$ 点有足够多次的导数，证明 $f(x)^n$ 和 x 轴至少有 $(n-1)$ 阶接触.

2. 设曲线 $y=f(x)$ 经过原点 O，并在原点与 x 轴相切. 证明在原点处曲线的曲率半径为

$$\rho=\lim_{x\to 0}\frac{x^2}{2y}.$$

*3. 设 K 是一个圆，这个圆在一点 P 处和一条给定的曲线相切，并通过曲线上一个相邻的点 θ. 证明：当 $\theta\to P$ 时圆 K 的极限就是这条曲线在 P 点的曲率圆.

*4. 证明：在曲率半径的极大值或极小值处，一曲线与其密切圆的接触的阶数至少是三.

*5. 证明：在曲率半径的极大值或极小值处，密切圆和曲线不会交叉，除非接触高于三阶.

*6. 求下列函数的最大值和最小值

(a) $\cos x\cosh x$, (b) $x+\cos x$.

*7. 试求函数 $y=e^{-\frac{1}{x^2}}$ 的最大值和最小值(参看第 272 页)

A.I.3 节，第 488 页

1. 若 f 在区间 $[0,1]$ 上连续，证明

$$\lim_{x\to 0} x\int_x^1\frac{f(z)}{z^2}dz=f(0).$$

2. 证明：函数 $y=(x^2)^x$, $y(0)=1$ 在 $x=0$ 点连续.

第六章 数 值 方 法

解一个分析问题总是不能臻于完善．虽然说解的存在性及某些基本性质的证明通常能令人满意，但是仍有有关的问题留下来待回答．譬如说，这解是用一个极限过程定义的，例如是用一个定积分定义的，于是实际地寻求这个极限的近似值并估计这些近似值的准确度的问题就提出来了．如果我们希望把分析方法应用于自然现象的描述和控制，而原则上它们又只能用近似的方式去描述，上述那样的问题就不仅在理论上带有根本的重要性，而且也是不可回避的．

因此，能给出解的数值解答并估计出它们所达到的准确度，这是一项很艰巨的任务．

近来，随着高速自动计算机的出现，"数值分析"在理论上和实践上都受到极大的推动． 它们出现于各种教科书里．[1]尽管如此，几百年来，许多大数学家，如牛顿，欧拉，特别是高斯，都对数值方法做出了很大的贡献．

在这一卷里，虽然我们不能对数值分析给予广泛的介绍，但是，至少将讨论某些简单的经典结果．

6.1 积分的计算

按照第二章的理论，虽然一个连续函数的积分的存在性是毫无疑义的了，但除少数情况外，这样一个积分的求值或"求面积"[2]

1) 例如，Hildebrand, Introduction to Numerical Analysis, McGraw-Hill Book Co., 1956; Householder, Principles of Numerical Analysis, McGraw-Hill Book Co., 1953; 和 Whittaker and Robinson, The Calculus of Observations, Blackie and Sons, Ltd., 1929.

2) "求面积"一词指出了求积步骤，也就是，通过寻找一个与之等积的正方形以测量曲线内的面积的步骤(例如"化圆为方"问题).

并不能通过初等函数来实现. 因此，我们必须寻求数值积分以及估计数值近似准确度的方法.

为了近似地计算积分

$$J=\int_a^b f(x)dx, \tag{1}$$

其中 $a<b$，我们用 $n+1$ 个点

$$x_\nu=a+\nu h,\quad nh=b-a\quad(\nu=0,1,\cdots,n) \tag{2}$$

将区间 $a\leqslant x\leqslant b$ 分成 n 等分，每分长为 $h=(b-a)/n$.
则

$$J=\sum_{\nu=1}^{n} J_\nu,$$

其中

$$J_\nu=\int_{x_{\nu-1}}^{x_\nu} f(x)dx; \tag{3}$$

这样，计算积分 J 的问题就归结为求面积 J_ν 的好的近似值的问题，而 J_ν 就是 J 所代表的整个面积被我们切割而成的宽为 h 的诸带形.

a. 矩形近似公式

来源于定积分的原始定义的最直接的近似公式给出关系式

$$J=\sum_{\nu=1}^{n} J_\nu\approx h(f_1+f_2+\cdots+f_n),$$

其中为了简便，我们记

$$f_\nu=f(x_\nu).$$

这里(以及整章)符号 $\approx$ 表示“近似相等”.

为了估计这个近似式的准确度或者“误差”，我们假设 $f(x)$ 在区间 $a\leqslant x\leqslant b$ 上连续，且导数一致有界: $|f'(x)|\leqslant M_1$. 那末，就容易证明(见第 534 页，6.1 节问题 4)

$$|J_\nu-hf_\nu|\leqslant\frac{M_1h^2}{2}, \tag{4}$$

因而

$$\left|J - h\sum_{\nu=1}^{n} f_\nu\right| \leqslant n\,\frac{M_1 h^2}{2} = \frac{1}{2}M_1(b-a)h. \tag{5}$$

于是,用有限和作定积分的近似的准确度,用第三章第 268 页的术语来说,是"网格宽度" h 阶的.

b. 改进的近似式——辛卜生法则

如果我们不是用矩形带而是用细长的梯形来近似面积 J_ν, 如图 6.1a 所示,那就不用费多少力气,而得到一个较好的近似式. 这近似公式(梯形公式)就是

$$J \approx \frac{1}{2}h(f_0 + f_1) + \frac{1}{2}h(f_1 + f_2) + \cdots + \frac{1}{2}h(f_{n-1} + f_n)$$

$$= h(f_1 + f_2 + \cdots + f_{n-1}) + \frac{h}{2}(f_0 + f_n), \tag{6}$$

因为在前式中,除首末两项外每个函数值出现二次.

一般比梯形公式稍微准确的近似式是,以子区间 $x_{\nu-1} \leqslant x \leqslant x_\nu$ 的中点 $x_{\nu-1} + \dfrac{h}{2}$ 处的曲线的切线为上边界的梯形去近似第 ν 个带形而得的近似式. 这个梯形的面积简单就是

$$hf_{\nu-1/2} = hf\left(x_{\nu-1} + \frac{h}{2}\right),$$

因而相加就得到切线公式

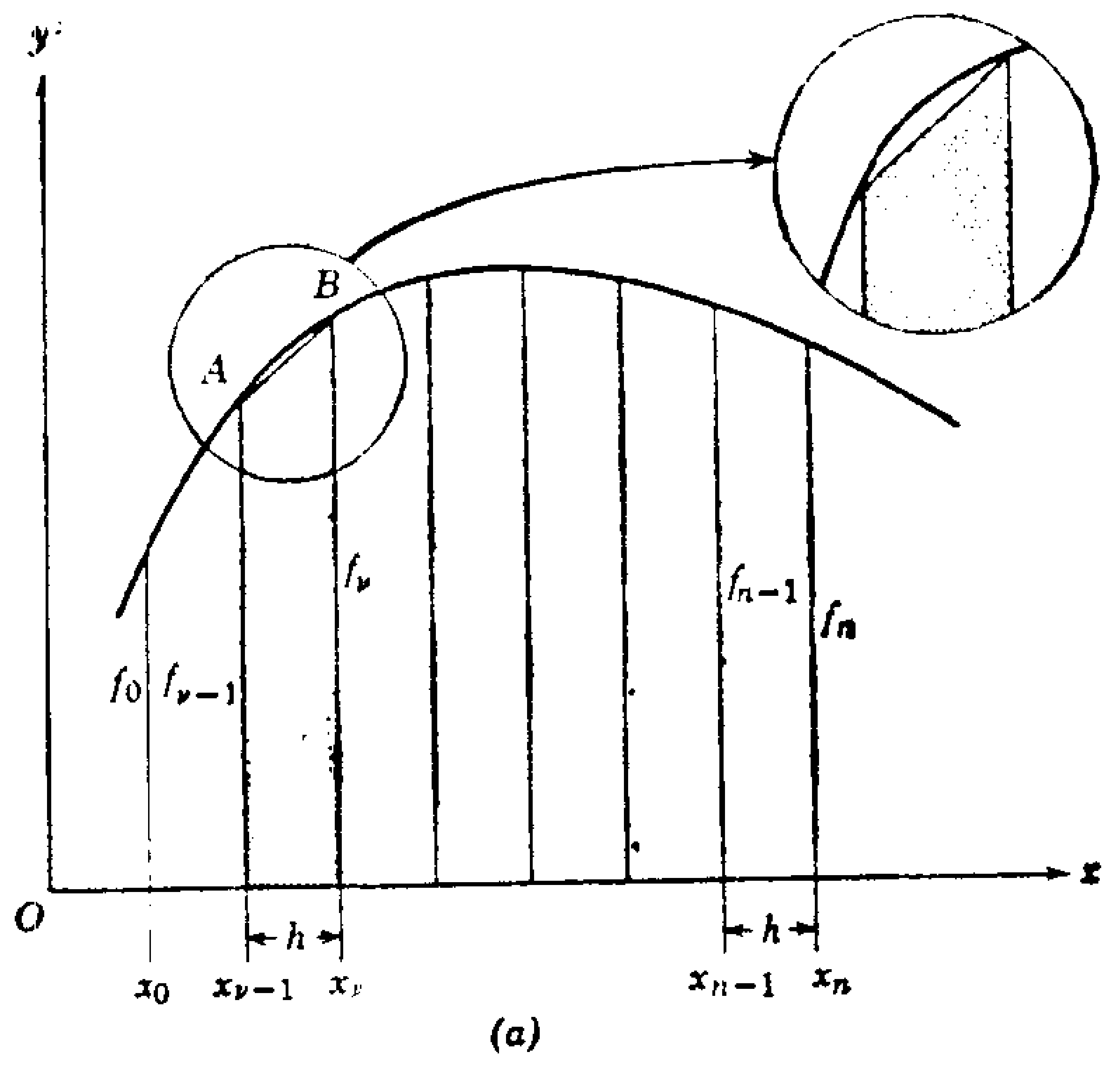
$$J \approx h(f_{1/2} + f_{3/2} + \cdots + f_{(2n-1)/2}). \tag{7}$$

正如我们将在第512页上看到的,当 f 的二阶导数在区间 $a \leqslant x \leqslant b$ 上连续,并存在某一个常数界 M_2, $|f''(x)| \leqslant M_2$, 那末这个近似式的准确度是 h^2 阶的.

最后,我们叙述著名的辛卜生(Simpson)近似式,它不用费多大劲,就给出一个准确得多的近似式,只要 f 的四阶导数存在并在区间内一致有界,即有一个常数 M_4 使得

$$|f^{(4)}(x)| \leqslant M_4.$$

对 $n = 2m$, 辛卜生公式就是



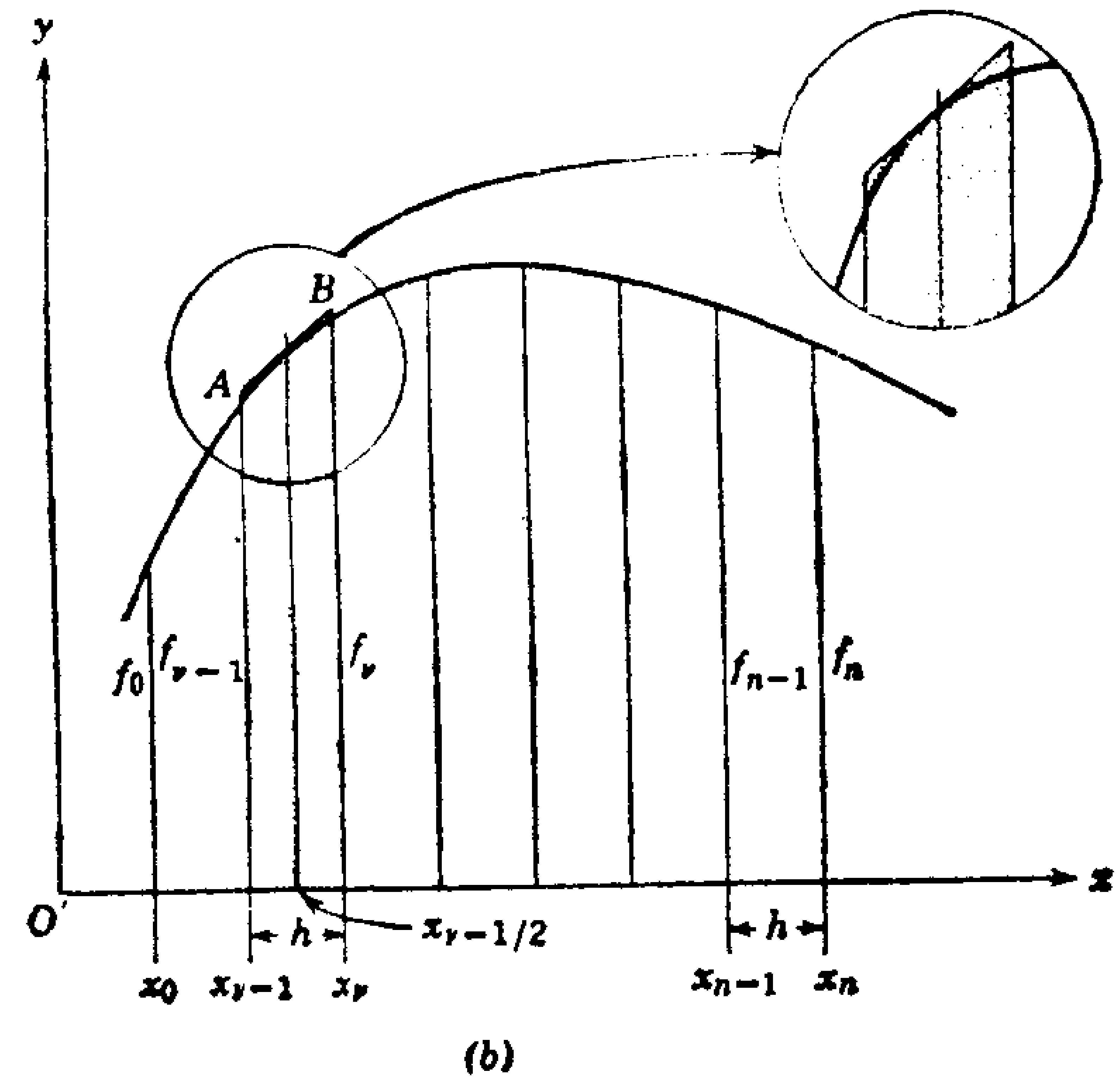


图 6.1 (a) 梯形公式 (b) 切线公式

$$J \approx \frac{4h}{3}(f_1 + f_3 + f_5 + \cdots + f_{2m-1}) + \frac{2h}{3}(f_2 + f_4 + f_6 + \cdots + f_{2m-2}) + \frac{h}{3}(f_0 + f_{2m}). \quad (8)$$

这公式容易求得，只需用一个宽为 $2h$ 的带，它的上界是在三个横

坐标为 $x_{\nu-1}$, $x_\nu=x_{\nu-1}+h$, 和 $x_{\nu+1}=x_{\nu-1}+2h$ 的点上与 f 重合的抛物线去近似第 ν 个及第 $\nu+1$ 个带所组成的区域(见图 6.2). 牛顿插值公式(第 498)给出了这个抛物线的方程:

$$y=f_{\nu-1}+(x-x_{\nu-1})\frac{f_\nu-f_{\nu-1}}{h}$$

$$+\frac{(x-x_{\nu-1})(x-x_{\nu-1}-h)}{2}\cdot\frac{f_{\nu+1}-2f_\nu+f_{\nu-1}}{h^2};$$

因此,我们有近似式

$$J_\nu+J_{\nu+1}\approx\int_{x_{\nu-1}}^{x_{\nu+1}}ydx=\int_{x_{\nu-1}}^{x_{\nu-1}+2h}ydx$$

$$=2hf_{\nu-1}+2h(f_\nu-f_{\nu-1})+\frac{\frac{8}{3}h-2h}{2}(f_{\nu+1}-2f_\nu+f_{\nu-1})$$

$$=\frac{h}{3}(f_{\nu-1}+4f_\nu+f_{\nu+1}).$$

现在,对 $\nu=1,3,5,\cdots,2m-1$, 把所有这些近似值相加或所有诸“带形对”的面积相加,就得到关于偶数 $n=2m$ 的上述公式.

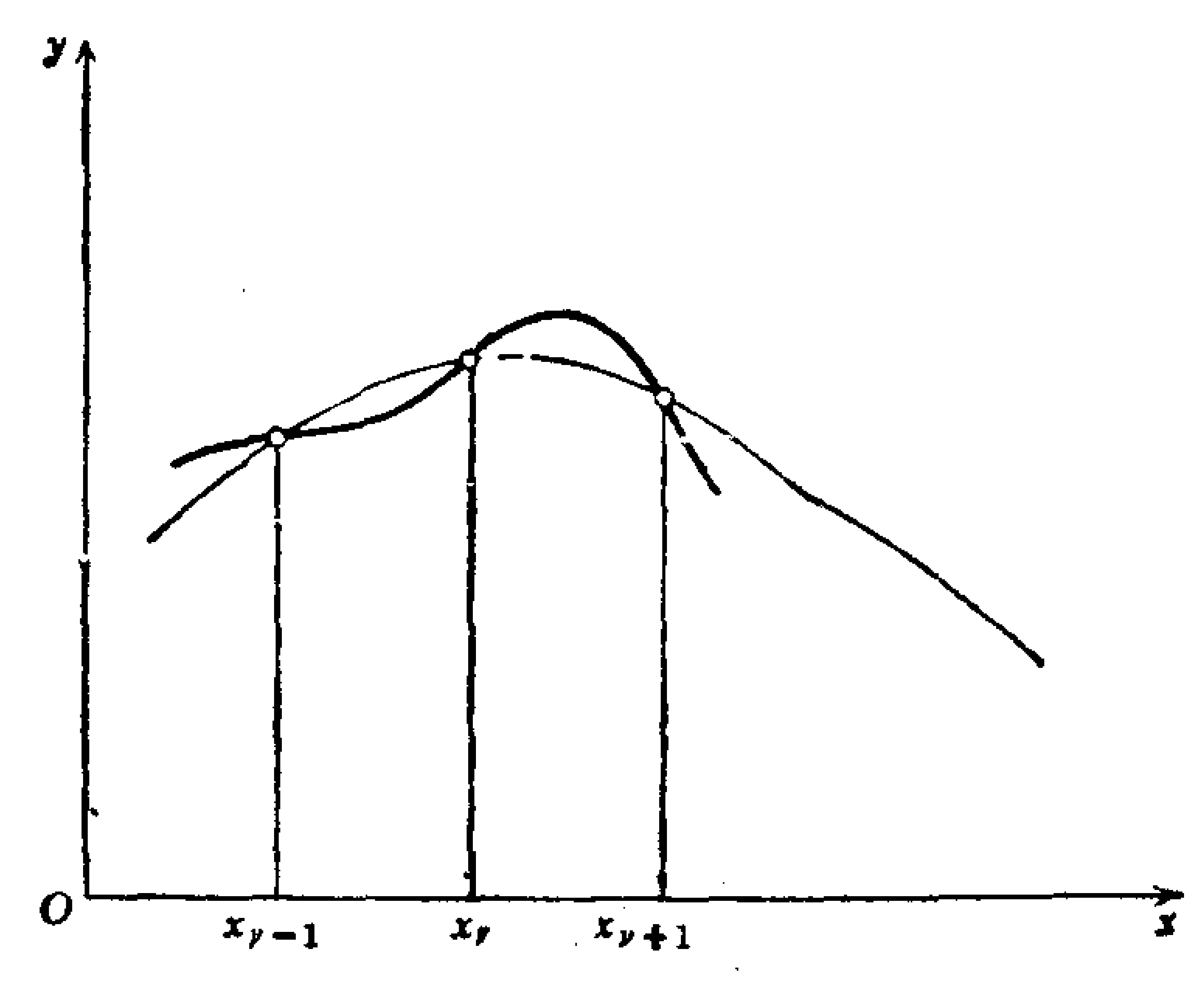


图 6.2　辛卜生法则

准确度

要估计我们的各近似公式的准确度并不难，因为求积的每一

步都是用一个容易积分的函数 $\phi(x)$（一个多项式）来逼近该区间上的函数 $f(x)$ 的．因此，积分公式的误差估计能够通过估计 $|f(x)-\phi(x)|$ 得到。

在切线公式里(第 508 页)，我们在区间 $[x_{\nu-1}, x_\nu]$ 上用中点 $x_\nu-(h/2)$ 处 $f(x)$ 的切线代替 $f(x)$，就是说，用

$$\phi(x)=f\left(x_\nu-\frac{h}{2}\right)+\left(x-x_\nu+\frac{h}{2}\right)f'\left(x_\nu-\frac{h}{2}\right)$$

代替 $f(x)$．根据拉格朗日余项形式的泰勒定理

$$f(x)=\phi(x)+\frac{1}{2}\left(x-x_\nu+\frac{h}{2}\right)^2 f''(\xi),$$

其中 ξ 位于 x 与 $x_\nu-\dfrac{h}{2}$ 之间．因此，对应于一条带形的误差由下式估计：

$$\begin{aligned}
|J_\nu-hf_{\nu-\frac{1}{2}}| &= \left|\int_{x_{\nu-1}}^{x_\nu}[f(x)-\phi(x)]dx\right| \\
&\leqslant \int_{x_{\nu-1}}^{x_\nu}|f(x)-\phi(x)|dx \\
&\leqslant M_2\int_{x_\nu-h}^{x_\nu}\frac{1}{2}\left(x-x_\nu+\frac{h}{2}\right)^2 dx \\
&= \frac{h^3}{24}M_2.
\end{aligned}$$

于是,对于切线公式里各区间所产生的总误差[1],我们求得上界为

$$n\frac{h^3}{24}M_2=\frac{h^2}{24}M_2(b-a).$$

作为估计误差的一个范例，我们把这个推导应用于其他求积公式．在梯形法则(6)中，我们在区间 $[x_{\nu-1}, x_\nu]$ 上用线性内插多项式

$$\phi(x)=f_{\nu-1}+(x-x_{\nu-1})\frac{f_\nu-f_{\nu-1}}{h}$$

1）这是使用近似公式固有的误差，叫做截断误差；实际上，因为计算中的舍入会产生附加误差．随着所取步数的增大（也就是 h 的缩小），舍入误差的总效果很可能会增大，而截断误差则在减小．

逼近 $f(x)$. 由 $n=1$ 时的插值公式余项的误差估计式 [见第 501 页等式(40)],我们得到

$$f(x)-\phi(x)=\frac{1}{2}(x-x_{\nu-1})(x-x_\nu)f''(\xi),$$

其中 ξ 位于 $x_{\nu-1}$ 与 x_ν 之间. 因此,计算 J_ν 时误差的绝对值至多是

$$M_2\int_{x_{\nu-1}}^{x_\nu}\left|\frac{1}{2}(x-x_{\nu-1})(x-x_\nu)\right|dx=\frac{h^3}{12}M_2,$$

因而,总的误差至多是这个量的 n 倍:

$$\frac{h^2}{12}M_2(b-a).$$

同样的技巧可用于辛卜生法则(8);把 $\phi(x)$ 取为一个在 $x_{\nu-1}$, x_ν, $x_{\nu+1}$ 三点处与 f 一致的二次多项式,就在 $J_\nu+J_{\nu+1}$ 内引出一个阶为 h^4 的误差. 而实际上,误差估计还可以改进一个数量级,只要取一个三次多项式 $\phi(x)$,使它在区间 $[x_{\nu-1},x_{\nu+1}]$ 上给出一个比二次多项式更好地近似于 f 的近似式,而且仍有相同的积分,就引导到关于积分 J 的同一个近似公式 (8). 我们简单地用 $x_{\nu-1}$, x_ν, $x_{\nu+1}$ 三点上与 $f(x)$ 重合的内插多项式,使之满足 $\phi'(x_\nu)=f'(x_\nu)$: 它的形式为

$$\phi(x)=A_0+A_1(x-x_{\nu-1})+A_2(x-x_{\nu-1})(x-x_\nu)$$
$$+A_3(x-x_{\nu-1})(x-x_\nu)(x-x_{\nu+1}).$$

此处,前三项代表在三点 $x_{\nu-1}$, x_ν, $x_{\nu+1}$ 上与 f 重合的二次插值多项式. 常数 A_3 必须由条件 $\phi'(x_\nu)=f'(x_\nu)$ 来决定.

末项

$$A_3(x-x_\nu+h)(x-x_\nu)(x-x_\nu-h)$$
$$=A_3[(x-x_\nu)^2-h^2](x-x_\nu)$$

显然是 $x-x_\nu$ 的一个奇函数,因此,从 $x_\nu-h$ 到 $x_\nu+h$ 的积分为零. 所以,对于 f 的近似式的误差,我们有估计[参看第 501 页(40), $n=3$, 而在 x_ν 点有二个重合的内插点]:

$$f-\phi=\frac{1}{4!}(x-x_{\nu-1})(x-x_\nu)^2(x-x_{\nu+1})f^{(4)}(\xi).$$

这就给出了，在计算 $J_\nu + J_{\nu+1}$ 时的误差估计，它是

$$\frac{h^5}{90}M_4,$$

因此，总误差估计是

$$\frac{n}{2}\frac{h^5}{90}M_4 = \frac{h^4}{180}(b-a)M_4.$$

自然地，在每个带形内用一个更高阶的多项式逼近函数 $f(x)$，我们就可以获得更高的准确度.

例. 我们用这些方法计算

$$\log_e 2 = \int_1^2 \frac{dx}{x}.$$

将区间 $1 \leqslant x \leqslant 2$ 分成长度为 $h = \frac{1}{10}$ 的十段，并且用梯形法则(6)，我们就得到.

$x_1 = 1.1$	$f_1 = 0.90909$
$x_2 = 1.2$	$f_2 = 0.83333$
$x_3 = 1.3$	$f_3 = 0.76923$
$x_4 = 1.4$	$f_4 = 0.71429$
$x_5 = 1.5$	$f_5 = 0.66667$
$x_6 = 1.6$	$f_6 = 0.62500$
$x_7 = 1.7$	$f_7 = 0.58824$
$x_8 = 1.8$	$f_8 = 0.55556$
$x_9 = 1.9$	$f_9 = 0.52632$
	和 6.18773
$x_0 = 1.0$	$\frac{1}{2}f_0 = 0.5$
$x_{10} = 2.0$	$\frac{1}{2}f_{10} = 0.25$
	$6.93773 \cdot \frac{1}{10}$

$$\log_e 2 \approx 0.69377.$$

因为被积函数的图形的凸的一边向着 x 轴，所以这个值太大了.

用切线法则 (7)，我们有

$$x_0 + \frac{1}{2}h = 1.05 \qquad f_{1/2} = 0.95238$$

$$x_1 + \frac{1}{2}h = 1.15 \qquad f_{3/2} = 0.86957$$

$$x_2 + \frac{1}{2}h = 1.25 \qquad f_{5/2} = 0.80000$$

$$x_3 + \frac{1}{2}h = 1.35 \qquad f_{7/2} = 0.74074$$

$$x_4 + \frac{1}{2}h = 1.45 \qquad f_{9/2} = 0.68966$$

$$x_5 + \frac{1}{2}h = 1.55 \qquad f_{11/2} = 0.64516$$

$$x_6 + \frac{1}{2}h = 1.65 \qquad f_{13/2} = 0.60606$$

$$x_7 + \frac{1}{2}h = 1.75 \qquad f_{15/2} = 0.57143$$

$$x_8 + \frac{1}{2}h = 1.85 \qquad f_{17/2} = 0.54054$$

$$x_9 + \frac{1}{2}h = 1.95 \qquad f_{19/2} = 0.51282$$

$$6.92863 \cdot \frac{1}{10}$$

$$\log_e 2 \approx 0.69284.$$

由于曲线的凸性，它又太小了.

对于同一个分割，用辛卜生法则 (8)，就得到一个准确得多的结果. 我们有

$$x_1 = 1.1 \qquad f_1 = 0.90909$$

$$x_3 = 1.3 \qquad f_3 = 0.76923$$

$$x_5 = 1.5 \qquad f_5 = 0.66667$$

$x_7 = 1.7$	$f_7 = 0.58824$
$x_9 = 1.9$	$f_9 = 0.52632$
	和　3.45955 · 4
	13.83820
$x_2 = 1.2$	$f_2 = 0.83333$
$x_4 = 1.4$	$f_4 = 0.71429$
$x_6 = 1.6$	$f_6 = 0.62500$
$x_8 = 1.8$	$f_8 = 0.55556$
	和　2.72818 · 2
	5.45636
	13.83820
$x_0 = 1.0$	$f_0 = 1.0$
$x_{10} = 2.0$	$f_{10} = 0.5$
	$20.79456 \cdot \frac{1}{30}$

$$\log_e 2 \approx 0.69315.$$

事实上

$$\log_e 2 = 0.693147\cdots.$$

6.2　数值方法的另一些例

a. "误差计算"

"误差计算"只是微分学的基本事实在数值上的应用：一个足够多次可微的函数 $f(x)$，在一点邻域内能够表示成一个线性函数带一个高于一阶的误差；表成一个二次函数带一个高于二阶的误差，等等.

考虑函数 $y = f(x)$ 的线性近似. 如果 $y + \Delta y = f(x + \Delta x) = f(x + h)$，根据泰勒定理，我们有

$$\Delta y = hf'(x) + \frac{h^2}{2} f''(\xi),$$

其中 $\xi = x + \theta h\,(0 < \theta < 1)$ 是一个无需准确知道的中间值. 如

果 $h = \Delta x$ 很小，我们就得到实用的近似式

$$\Delta y \approx hf'(x).$$

这样，我们就用导数代替近似等于其值的差商，并且用近似相等的 h 的线性式代替 y 的增量.

这个简单事实以下述方式用于数值计算. 设两个物理量 x 与 y 满足关系 $y = f(x)$. 那么我们要问: 由于 x 测量的不准确，对确定 y 有什么影响? 如果我们用不准确值 $x + h$ 代换"真"值 x，那么 y 的相应值与真值 $y = f(x)$ 之差的大小就等于 $\Delta y = f(x + h) - f(x)$. 因此，这误差可用上述关系近似地给出.

我们举例说明这种线性近似的用法.

例. (a) 在 ΔABC 中(参看图 6.3)，假设准确地测得了边 b 和 c，而角 $\alpha = x$ 的测量则在误差范围 $|\Delta x| < \delta$ 之内. 试问第三边的值 $y = a = \sqrt{b^2 + c^2 - 2bc\cos\alpha}$ 的相应误差是多少?

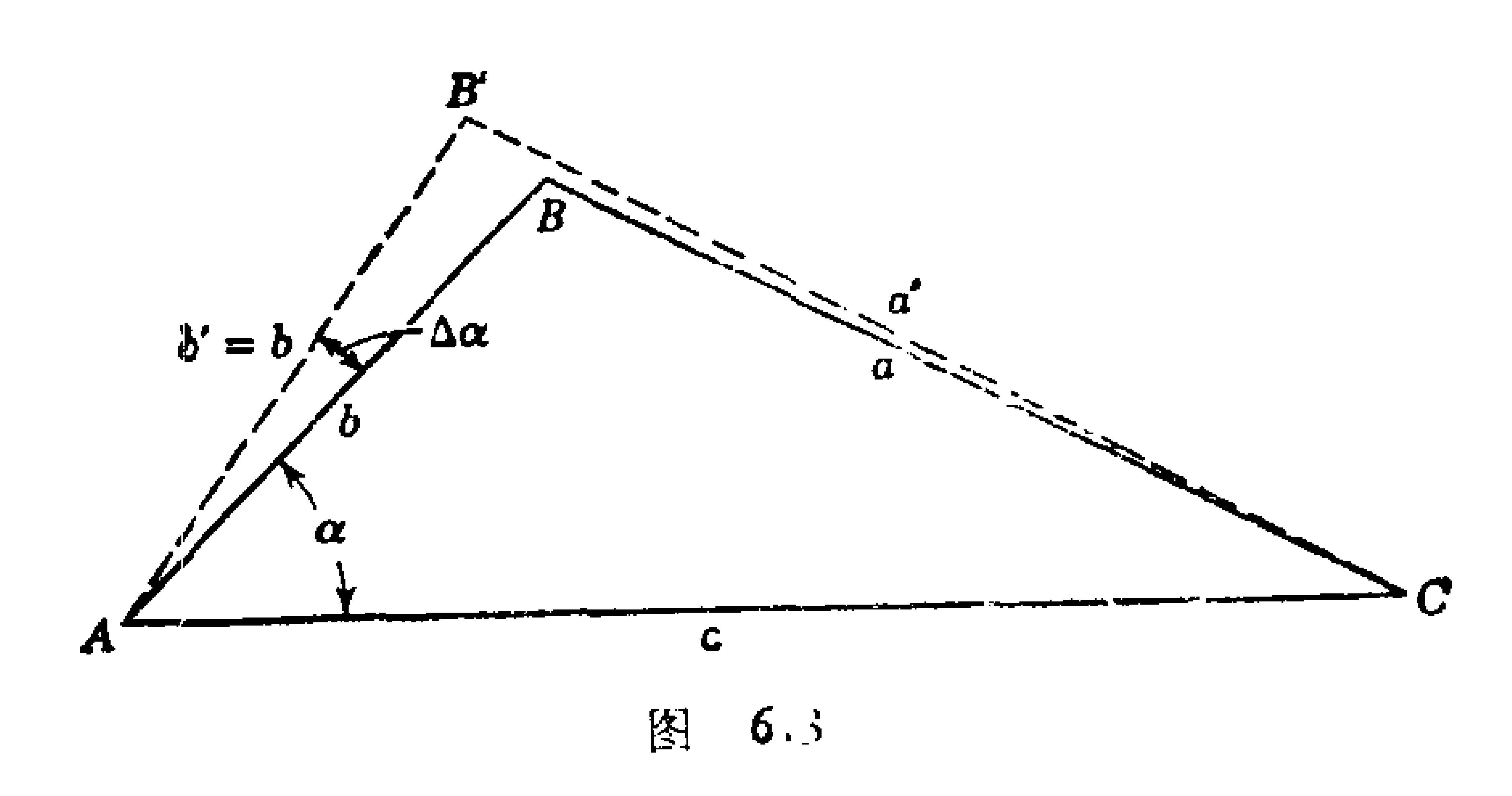


图 6.3

我们有 $\Delta a \approx (bc\sin\alpha\,\Delta\alpha)/a$; 因此，百分误差是

$$\frac{100\Delta a}{a} \approx \frac{100bc}{a^2}\sin\alpha\,\Delta\alpha.$$

特别地，当 $b = 400$ 米，$c = 500$ 米和 $\alpha = 60°$ 时，我们有 $y = a = 458.2576$ 米，所以

$$\Delta a \approx \frac{200000}{458.2576} \times \frac{1}{2}\sqrt{3}\,\Delta\alpha.$$

如果 $\Delta\alpha$ 能在弧为 10 秒之内测得，就是说，如果

$$\Delta\alpha = 10'' = 4846 \times 10^{-8} \text{ 弧度},$$

我们求得最坏结果也就是

$$\Delta a \approx 1.83 \text{ 厘米};$$

因此，误差至多约为 0.004%。

(b) 下面这个例说明了线性化在物理问题中的用法。

由试验得知：如果温度为 t_0 时一个金属棒的长度为 l_0，则温度为 t 时，它的长度将是 $l = l_0(1 + \alpha(t - t_0))$；其中 α 仅仅依赖于 t_0 和棒的构成原料。现在，如果一个摆钟在温度为 t_0 时保持准时，试问当温度升到 t_1 时，它每天将要慢多少秒？

关于振动周期 $T(l)$，我们有

$$T(l) = 2\pi\sqrt{\frac{l}{g}}.$$

(见第 435 页)；因此，

$$\frac{dT}{dl} = \frac{\pi}{\sqrt{lg}}.$$

如果长度变化为 Δl，振动周期的相应变化就是

$$\Delta T \approx \frac{\pi \Delta l}{\sqrt{l_0 g}},$$

其中 $l_1 = l_0(1 + \alpha(t_1 - t_0))$，$\Delta l = \alpha l_0(t_1 - t_0)$。这就是每次振动失去的时间。一秒钟将慢 $\dfrac{\Delta T}{T} \approx \Delta l/2l_0$；因此，这个摆钟每天要慢 $43200\ \Delta l/l_0 = 43200\alpha(t_1 - t_0)$ 秒。

在这个例中以及许多所考虑的函数是几个因子的积的其他情况中，我们在微分之前，对等式两边取对数，可以简化计算。在这个例中，我们有

$$\log T = \log 2\pi - \frac{1}{2}\log g + \frac{1}{2}\log l;$$

微分得，

$$\frac{1}{T}\frac{dT}{dl} = \frac{1}{2l}.$$

用 $\Delta T/\Delta l$ 代换 $\dfrac{dT}{dl}$，给出

$$\frac{\Delta T}{T} \approx \frac{\Delta l}{2l},$$

与前面的结果一致.

*b. π 的计算

另一个与前面不同的例是使用了特别技巧的经典例子，尽管也许由于有了现代计算技术这些技巧已经不用了.

用反正切级数的莱布尼兹级数 $\frac{\pi}{4}=1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\cdots$ [第 468 页第 5.2 节 (7)]计算 π 并不适用，因为它收敛得太慢. 然而使用下述技巧，我们可以较容易地计算 π. 如果在正切的加法定理

$$\tan(\alpha+\beta)=\frac{\tan\alpha+\tan\beta}{1-\tan\alpha\tan\beta}$$

中，引入反函数 $\alpha=\arctan u$，$\beta=\arctan v$，我们就得到公式

$$\arctan u+\arctan v=\arctan\left(\frac{u+v}{1-uv}\right).$$

现在，选择 u 和 v，使得 $(u+v)/(1-uv)=1$，我们在右边得到值 $\frac{\pi}{4}$; 而如果 u 与 v 是小的数，我们借助于已知级数就可以容易地算出左边. 例如象欧拉做过的，如果令 $u=\frac{1}{2}$，$v=\frac{1}{3}$，我们就得到

$$\frac{\pi}{4}=\arctan\frac{1}{2}+\arctan\frac{1}{3}. \tag{9}$$

如果再注意到 $\dfrac{\frac{1}{3}+\frac{1}{7}}{1-\frac{1}{21}}=\frac{1}{2}$，我们有 $\arctan\frac{1}{2}=\arctan\frac{1}{3}+\arctan\frac{1}{7}$，根据 (9)，就有

$$\frac{\pi}{4}=2\arctan\frac{1}{3}+\arctan\frac{1}{7}.$$

利用这个公式，维加（Vega）将数 π 的值计算到 140 位.

借助等式 $\left(\frac{1}{5}+\frac{1}{8}\right)\Big/\left(1-\frac{1}{40}\right)=\frac{1}{3}$，我们进一步得到

$$\operatorname{arc\,tan}\frac{1}{3}=\operatorname{arc\,tan}\frac{1}{5}+\operatorname{arc\,tan}\frac{1}{8},$$

或者

$$\frac{\pi}{4}=2\operatorname{arc\,tan}\frac{1}{5}+\operatorname{arc\,tan}\frac{1}{7}+2\operatorname{arc\,tan}\frac{1}{8}.$$

这个展式对于借助级数 $\operatorname{arc\,tan}x=x-\frac{x^3}{3}+\frac{x^5}{5}-\cdots$ 计算 π 是非常有用的；因为如果我们用值 $\frac{1}{5}$，$\frac{1}{7}$ 或 $\frac{1}{8}$ 代替 x，则由于项减小得很快，只用很少几项，我们就可得到一个高准确度的结果.

对这些技巧，以及这些艺术处理不特别感兴趣的读者，对原理有个理解就可以了.

*c. 对数的计算

为了求对数的数值，我们来变换对数级数[第 467 页，(5)]

$$\frac{1}{2}\log\frac{1+x}{1-x}=x+\frac{x^3}{3}+\frac{x^5}{5}+\cdots,$$

其中 $0<x<1$. 将

$$\frac{1+x}{1-x}=\frac{p^2}{p^2-1},\qquad x=\frac{1}{2p^2-1}$$

代入级数得

$$\log p=\frac{1}{2}\log(p-1)+\frac{1}{2}\log(p+1)+\frac{1}{2p^2-1}+\frac{1}{3(2p^2-1)^3}+\cdots,$$

其中 $2p^2-1>1$ 或 $p^2>1$. 如果 p 是一个整数而且 $p+1$ 可以分解成较小的整数因子（譬如，如果 $p+1$ 是偶数），后面这个级数就把 p 的对数表示成较小整数的对数加一个级数，而级数的诸项

减小得很快，因而，它的和只用很少几项就可以充分准确地计算出来. 因此，只要我们已经算出了 log 2 的值（例如，可用第 513 页上的积分表示法进行计算），我们就能从这个级数逐步算出任何素数的对数值，进而也能算出任何整数的对数值.

这样确定的 $\log p$ 的准确度可用几何级数来估计，这比用余项的一般公式容易得多. 关于级数的余项 R_n，也就是 $1/n(2p^2-1)^n$ 这一项之后的所有项的和，我们有

$$R_n<\frac{1}{(n+2)(2p^2-1)^{n+2}}\left[1+\frac{1}{(2p^2-1)^2}+\frac{1}{(2p^2-1)^4}+\cdots\right]$$

$$=\frac{1}{(n+2)(2p^2-1)^n}\cdot\frac{1}{(2p^2-1)^2-1},$$

而这个公式立即给出了所要求的误差估计.

作为举例，让我们用级数的前四项计算 $\log_e 7$（在 $\log_e 2$ 和 $\log_e 3$ 已经求出了数值的假设之下）. 我们有

$$p=7,\qquad 2p^2-1=97,$$

$$\log 7=2\log 2+\frac{1}{2}\log 3+\frac{1}{97}+\frac{1}{3\cdot 97^3}+\cdots$$

$$\frac{1}{97}\approx 0.01030928,\quad \frac{1}{3\cdot 97^3}\approx 0.00000037,$$

$$2\log 2\approx 1.38629436,\quad \frac{1}{2}\log 3\approx 0.54930614;$$

因此

$$\log_e 7\approx 1.94591015.$$

估计误差给出

$$R_n<\frac{1}{5\cdot 97^3}\times\frac{1}{97^2-1}<\frac{1}{36\times 10^9}.$$

但是，我们注意到，我们相加的四个数中每一个数都只给出范围在 5×10^{-9} 之内的误差. 这就使得 log 7 的计算值的末位可能差 2. 然而，事实上，末位也是准确的.

6.3 方程的数值解法

关于方程 $f(x)=0$ 的数值解法我们给一些注释，此处 $f(x)$

不需要是一个多项式.[1] 我们从诸根之一的试验性的第一个值 x_0 开始,然后改进这个近似值.根的第一个近似值如何选取和该近似式有多好,可以暂不考虑. 例如,对第一个近似值,我们可以做个粗略的猜想,或者更好些,从函数 $y=f(x)$ 的图象上取,因为函数图象与 x 轴的交点就表示了所要求的根.

然后,我们试着用一种程序或一个映射去改进这个近似值,从而把值 x_0 变成"第二个近似值",并且重复这种程序. 数值地解方程 $f(x)=0$ 就是重复上述的逐次逼近(或者,如人们所说"迭代"程序),以期望迭代值 $x_1, x_2, \cdots, x_n$ 满意地收敛到根 ξ.我们将考虑各种这样的程序,并简短地讨论它们的准确度.

a. 牛顿法

方法综述. 牛顿迭代程序基于微分学的基本原理——在靠近切点的邻域内,用切线代替曲线. 从方程 $f(x)=0$ 的根 ξ 的第一个近似值 x_0 出发,我们考虑函数 $y=f(x)$ 图形上的点,其坐标是 $x=x_0,\ \ y=f(x_0)$. 为了找出曲线与 x 轴的交点 ξ 的一个较好的

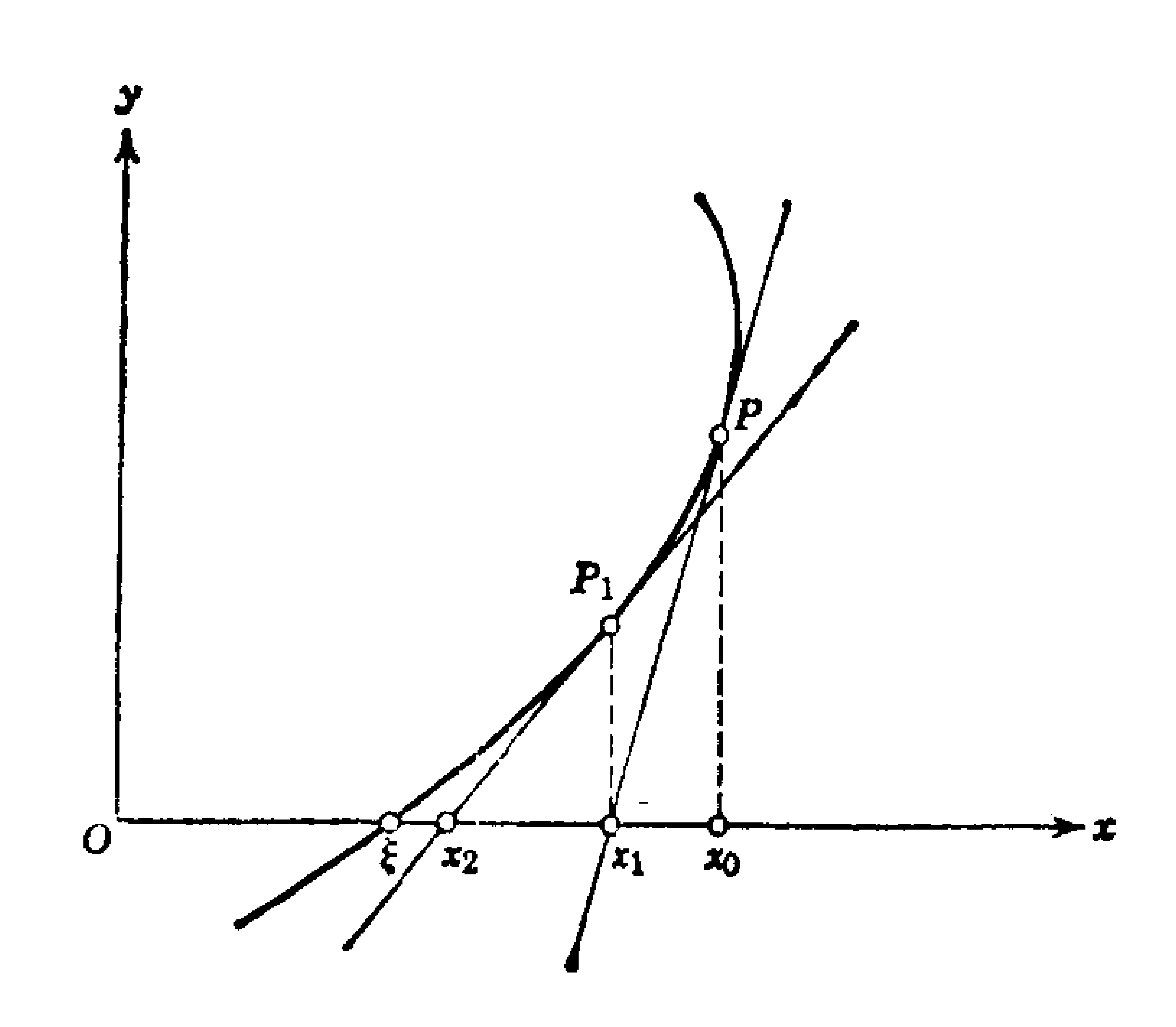


图 6.4 近似公式的牛顿方法

---

1) 当然,我们关心的仅仅是确定 $f(x)=0$ 的实数根.

近似值，我们取定点 x_1，它是在点 $x=x_0$，$y=f(x_0)$ 处的切线与 x 轴的交点．这个交点的横坐标 x_1 代表了一个新的，而且在某些情况下是一个比 x_0 更逼近所求方程根 ξ 的近似值．

图 6.4 立刻给出

$$\frac{f(x_0)}{x_0-x_1}=f'(x_0);$$

因此，新的近似值

$$x_1=x_0-\frac{f(x_0)}{f'(x_0)}. \tag{10}$$

现在从 x_1 作为一个近似值出发，我们重复这程序，去求 $x_2=x_1-f(x_1)/f'(x_1)$，然后，如此继续下去．

这种程序之有用，本质上依赖于曲线 $y=f(x)$ 的性质．在图 6.4 指出的情况中，逐次近似值 x_n 以越来越高的精确度收敛到所求的根 ξ．

然而，图 6.5 指出了一个原始值 x_0 的似是而非的选择，我们的作图根本不收敛到所求的根．因此，必需一般地考察在怎样的情况下，牛顿法给出方程解的有用的近似值．

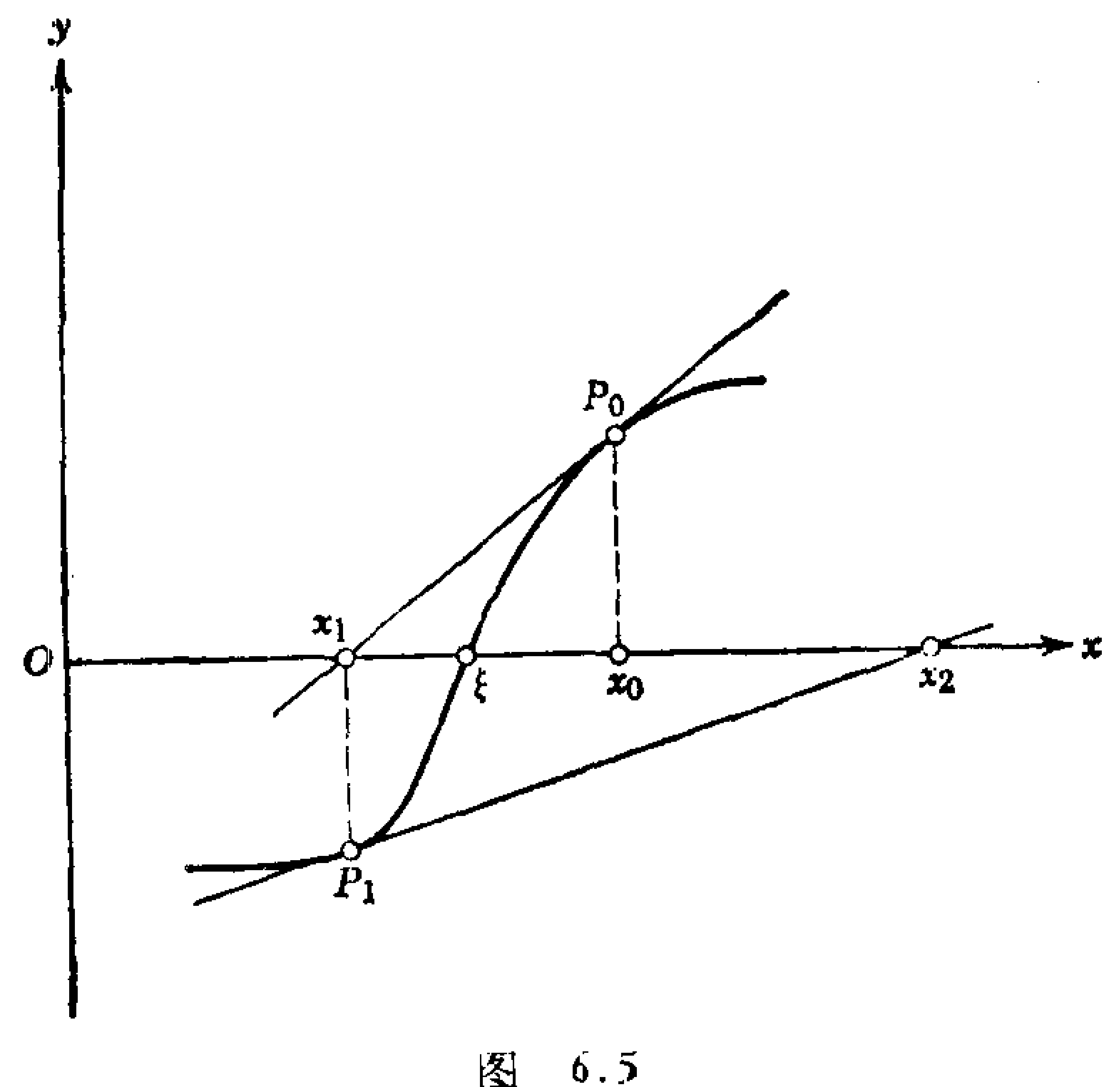


图 6.5

牛顿法的二次收敛

假设在根 ξ 附近一个足够宽的区间里，二阶导数 $f''(x)$ 不"太大"，而一阶导数 $f'(x)$ 不"太小"，则牛顿近似法的主要之点是逐次"误差"

$$h_1 = \xi - x_1,\quad h_2 = \xi - x_2, \cdots \quad h_n = \xi - x_n, \cdots$$

在意义 $|h_{n+1}| \leqslant \mu h_n^2$ 之下二次收敛到零，其中 μ 是一个固定的常数. 这指出了一个极快的收敛速度. 如果我们把这个不等式写成 $|h_{n+1}\mu| \leqslant |h_n\mu|^2$ 的形式，它就意味着，譬如当 $|h_n\mu| < 10^{-m}$ 时，我们有 $|h_{n+1}\mu| < 10^{-2m}$，即 μx_n 中"有效位"的位数是每步成倍增加的.

这个二次收敛的证明是直截了当的. 从关系式 $x_{n+1} = x_n - f(x_n)/f'(x_n)$ 和 $f(\xi) = 0$，我们得到

$$h_{n+1} = \xi - x_{n+1} = \xi - x_n - \frac{f(\xi) - f(x_n)}{f'(x_n)}.$$

根据泰勒公式

$$f(\xi) - f(x_n) = (\xi - x_n)f'(x_n) + \frac{1}{2}(\xi - x_n)^2 f''(\eta),$$

其中 η 位于 ξ 与 x_n 之间. 因此，

$$h_{n+1} = -\frac{f''(\eta)}{2f'(x_n)} h_n^2. \tag{11}$$

为了确立收敛性，我们设 x_n 已经属于一个固定区间 $\xi - \delta < x < \xi + \delta$，在此区间内 $|f''|$ 有最大值 M_2，$|f'|$ 有正的最小值 m_1，并且 δ 很小，使得 $\frac{1}{2}\delta M_2/m_1 < 1$. 令 $\mu = \frac{1}{2}M_2/m_1$，我们有 $\mu\delta < 1$ 以及

$$|h_{n+1}| \leqslant \mu |h_n|^2 \leqslant \mu\delta |h_n| < |h_n|.$$

这个不等式首先指出，x_{n+1} 仍属于同一个 ξ 的 δ 邻域，因此，这个结论可重复使用. 所以，只要 x_0 位于 ξ 的 δ 邻域内，那么所有后续的 x_n 也将属于同一个 ξ 的 δ 邻域. 因为从 $|h_{n+1}| \leqslant \mu\delta |h_n|$，可推出 $|h_{n+1}| \leqslant (\mu\delta)^{n+1}|h_0|$，这就意味着 $h_n \to 0$ 或 $x_n \to \xi$；另外，递减的二次规律 $|h_{n+1}| \leqslant \mu |h_n|^2$ 对误差也成立. 于是很清

楚，牛顿法将为我们提供一个一定收敛到解 ξ 的序列 x_n，只要假定 f' 与 f'' 存在并在 ξ 附近连续，$f'(\xi) \neq 0$。而且 x_0 已足够接近 ξ。近似式的二次性质通常是牛顿法超越其他方法的一个决定性的优点(见第 529 页)。

*b. 假位法

牛顿法是一个比较古老的方法——假位法——的极限情况。在这个比较古老的方法中，割线代替了切线。假设在所求的曲线与 x 轴的交点的邻域内，我们已知两个点 (x_0, y_0) 和 (x_1, y_1)。如果我们用连结这两点的割线代替曲线，这条割线与 x 轴的交点就能够作为一个所求的方程根[1]的一个改进了的近似值。对于交点的横坐标 ξ，我们有 (图 6.6)

$$\frac{\xi - x_0}{f(x_0)} = \frac{\xi - x_1}{f(x_1)}, \tag{12}$$

由此导出

$$\xi = \frac{x_0 f(x_1) - x_1 f(x_0)}{f(x_1) - f(x_0)}$$

$$= \frac{x_0 f(x_1) - x_0 f(x_0) + x_0 f(x_0) - x_1 f(x_0)}{f(x_1) - f(x_0)}$$

或

$$\xi = x_0 - \frac{f(x_0)}{\dfrac{f(x_1) - f(x_0)}{x_1 - x_0}} \tag{13}$$

这个从 x_0 和 x_1 确定进一步近似值 ξ 的公式构成了假位法。如果函数的一个值为正，而另一个为负时，譬如说象图 6.6 中那样，$y_0 > 0$，$y_1 < 0$，那末假位法就更显得有用。

牛顿的近似公式是作为 $x_1 \to x_0$ 时的一种极限情况，因为当 x_1 趋向 x_0 时，公式 (13) 右边第二项的分母趋向 $f'(x_0)$。

虽然可以认为假位法比牛顿法更基本，但因为牛顿法只需要

---

1) 这实质上相当于线性内插用于反函数。

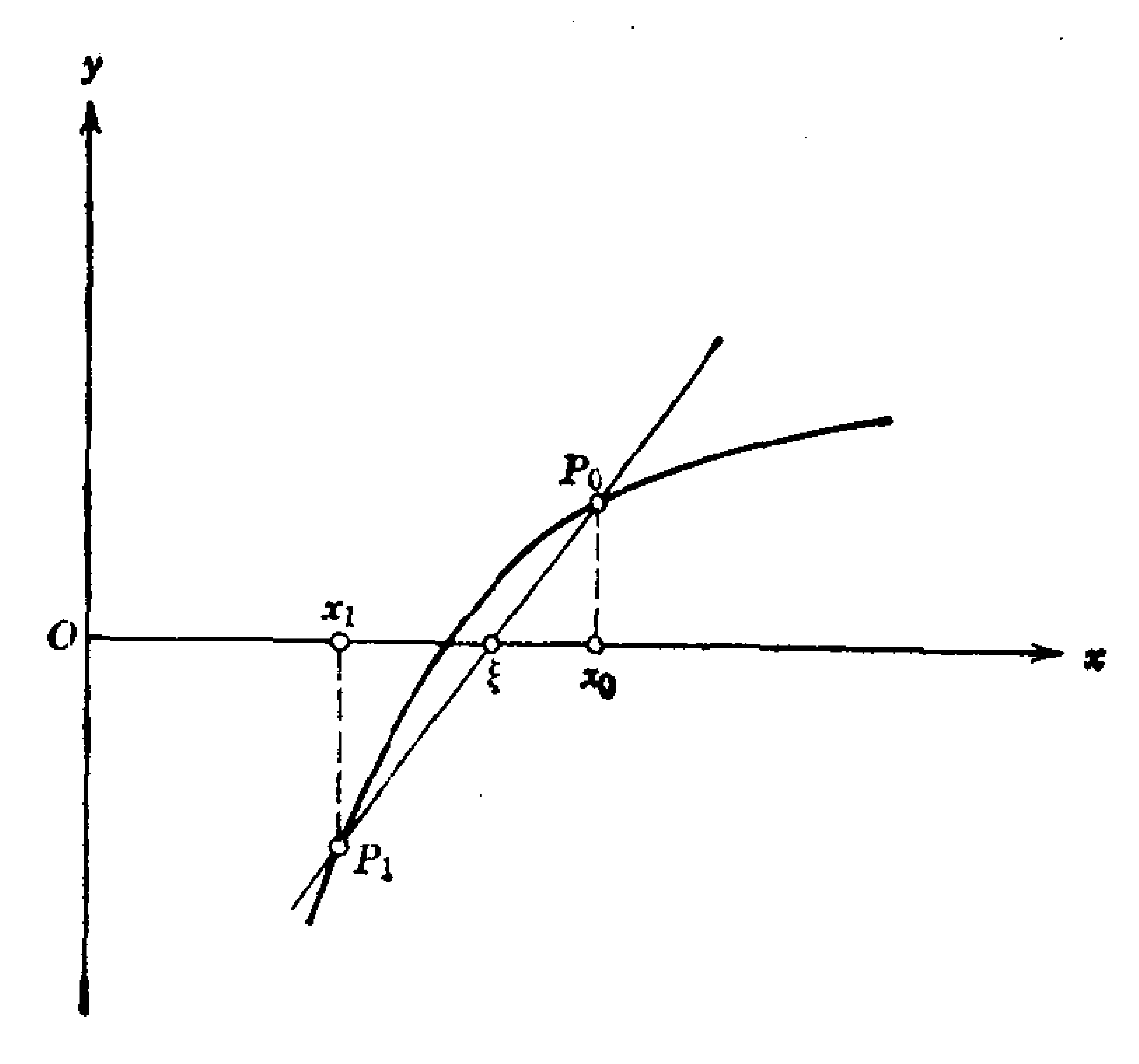


图 6.6 假位法

一个 x 值作为初始近似值，而不是两个值，所以牛顿法有很大的方便性。

c. 迭代法

迭代模式．现在我们转向一个应用极广的模式来解

$$x = \phi(x)$$

这种形式的方程，其中 ϕ 是一个具有连续导数的连续函数．如果我们令 $\phi(x) = x - c(x)f(x)$，这里 $c(x)$ 是任一不为 0 的函数，那末，解形如 $f(x) = 0$ 的方程就可简化为解形如 $x = \phi(x)$ 的方程．

在这个特别有启发性的迭代法[1]里，我们也从适当选择初始近似值 x_0 开始，从而由条件

$$x_{n+1} = \phi(x_n), \qquad n = 0, 1, 2, \cdots,$$

确定一个序列 $x_1, x_2, x_3, \cdots$．如果这个"迭代"序列 x_n 收敛到一个极限 ξ，那么 $\xi = \phi(\xi)$ 就是我们方程的一个解，因为这时

1）有时叫做逐次逼近法．为了解这类或那类方程，这个方法被用于许多不同的数学内容．

$\lim\limits_{n\to\infty} x_{n+1} = \xi$ 并且由于函数 ϕ 的连续性而有 $\lim\limits_{x\to\infty} \phi(x_n) = \phi(\xi)$.

收敛性. 序列值 x_n 在迭代程序中收敛到一个解，这是在一个非常普遍的条件下成立的：如果第一个近似值 x_0 位于围绕着解 ξ 的一个区间 J 内[1)]，而在这区间内有

$$|\phi'(x)| < q,$$

其中 $q < 1$ 是一个常数，则 x_n 收敛到 ξ.

因为设 x_0 位于 J 内，我们就有

$$x_1 - \xi = \phi(x_0) - \phi(\xi).$$

根据中值定理，这个方程的右边等于 $(x_0 - \xi)\phi'(\bar{x})$，其中 $\bar{x}$ 位于 J 内. 于是，根据我们的假设就有

$$|x_1 - \xi| \leqslant q|x_0 - \xi|,$$

所以，x_1 属于 J，因而又有

$$|x_2 - \xi| \leqslant q|x_1 - \xi| \leqslant q^2|x_0 - \xi|.$$

一般地，我们得到

$$|x_n - \xi| \leqslant q^n|x_0 - \xi|;$$

因为当 $n \to \infty$ 时 $q^n \to 0$，所以我们的论断得证.

另外，如上所见，当在 ξ 附近的区间内 $\phi'(x) > 1$ 时，迭代序列 x_n 不收敛；如果 $|\phi'(\xi)| = 1$，我们不能做出一般结论.

吸引性和排斥性不动点

通过映射或变换来研究迭代程序是很有用的. 函数 $y=\phi(x)$ 代表一个变换，它把数轴上的一个点 x 映射到这个数轴上的一个象点 y (见第 20 页). 因此，解 ξ 就是一个不被变换 ϕ 所改变的点，即所谓的不动点，于是，问题就是找映射的不动点；如已看到的，当 $|\phi'(\xi)| \leqslant q < 1$ 时，这个问题可用迭代法解决.

在 $|\phi'(x)| < q < 1$ 的情况下，根或不动点 ξ 的邻域的映射 $y = \phi(x)$ 具有收缩性质，就是它缩短了原来点到不动点的距离. 这种收缩映射的不动点称为吸引性不动点. 它们的迭代结构象一个公比为 q 的几何级数的项一样收敛.

---

1) 虽然 ξ 是未知的，但是，通常我们可以事先定出这样一个区间.

如果根 ξ 或我们的变换的相应不动点在一个区间内有 $|\phi'(x)| > r$，其中 r 是一个大于1的常数，则变换是扩张的，迭代程序发散，而不动点称为排斥性的。

如果在不动点上，我们有 $|\phi'(\xi)| = 1$，就不能做出关于迭代收敛性的一般结论。这样的不动点有时称为中立的(中性的)。

下述意见应该强调：映射 ϕ 的一个不动点 ξ 也自然而然地是逆映射的一个不动点：$\xi = \psi(\xi)$。如果在根 ξ 的一邻域内 $|\phi'(\xi)| > 1$，而且 $x = \psi(y)$ 是 ϕ 的反函数，则 $|\psi'(\xi)| < 1$。于是 ξ 就是这个逆映射的一个吸引性不动点，因而可能用对于逆映射是收敛的迭代模式代替原来发散的迭代模式。作为一个例，我们考虑方程

$$x = \tan x.$$

从函数 $y = x$ 和 $y = \tan x$ 的图象可清楚地看到，它们在区间

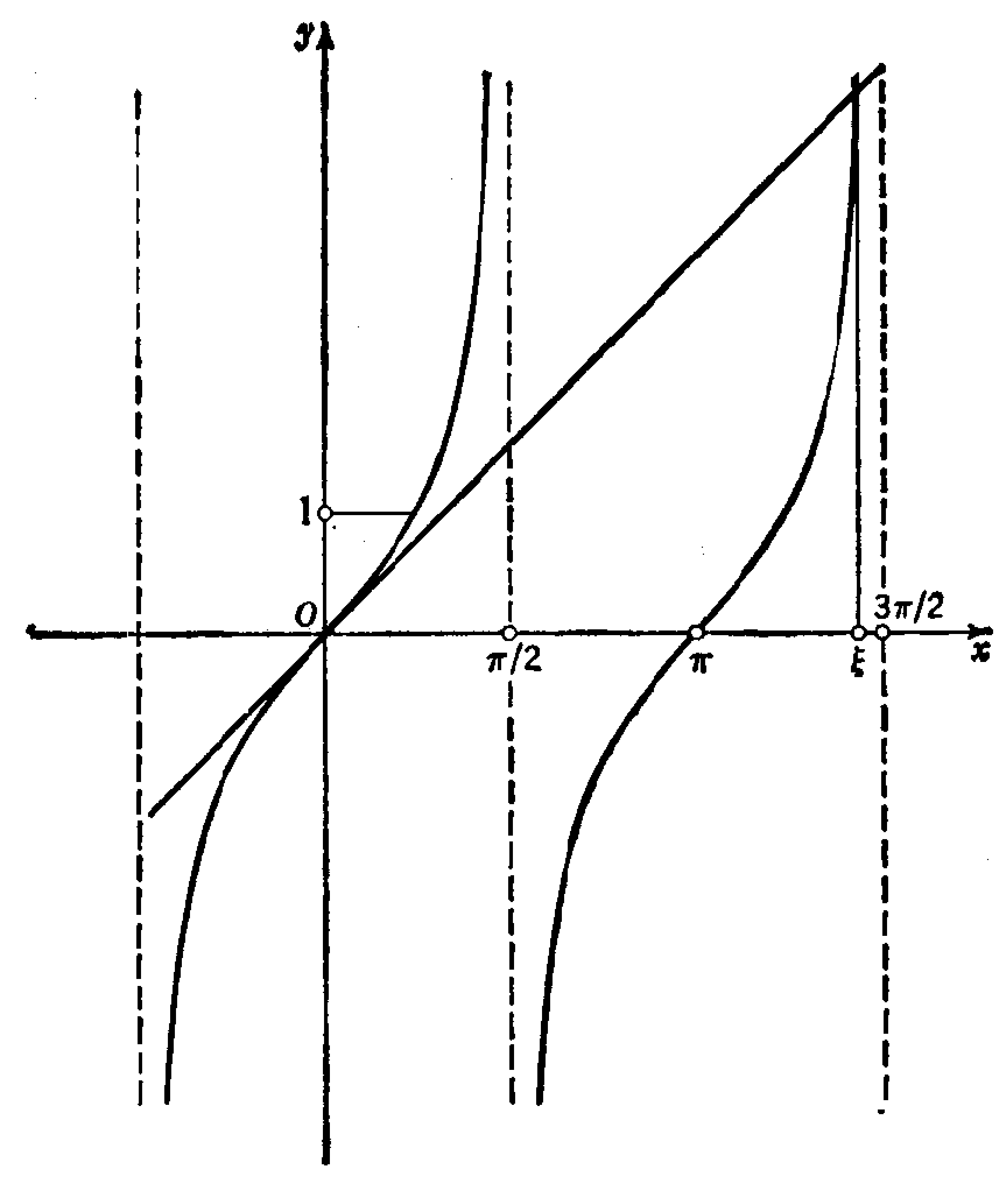


图 6.7 曲线 $y = \tan x$ 和 $y = x$ 的交点 (ξ, ξ)

$\pi < x < \frac{3}{2}\pi$ 内交于某点，因此，我们的方程也将在该区间有一个根 ξ. 因为

$$\frac{d\tan x}{dx} = \frac{1}{\cos^2 x} > 1,$$

所以，以区间内的任何点 x_0 为初始值的迭代程序都不会收敛. 然而，如果我们把方程写成逆形式(用符号 arc tan x 记主值分支)

$$x = \arctan x + \pi,$$

我们就得到一个收敛的迭代序列. 因为这里

$$\frac{d}{dx}\arctan x = \frac{1}{1+x^2} < 1,$$

所以用 $x_{n+1} = \arctan x_n + \pi$ 以及，譬如说，$x_0 = \pi$ 所定义的序列收敛到 ξ.

d. 迭代与牛顿程序

如前所述，解形式为 $f(x) = 0$ 的方程可以归结为解形式为 $x = \phi(x)$ 的方程，只要我们把 $\phi(x)$ 选成形式为

$$\phi(x) = x - c(x)f(x).$$

的任何一个式子，其中 $c(x)$ 是一个非 0 函数. 如果我们想用迭代法解所得的方程 $x = \phi(x)$，我们就必须通过适当选取 $c(x)$ 保证映射 ϕ 的不动点 ξ 是“吸引的”，也就是说，必须有 $|\phi'(\xi)| < 1$. 现在，对于 $f(\xi) = 0$ 的解 ξ，我们有

$$\phi'(\xi) = 1 - c'(\xi)f(\xi) - c(\xi)f'(\xi) = 1 - c(\xi)f'(\xi).$$

最简单的选法是取 $c(x)$ 为 $\frac{1}{f'(x)}$. 于是当然有 $|\phi'(\xi)| = 0 < 1$. $c(x)$ 的这种选法导致迭代序列

$$x_{n+1} = \phi(x_n) = x_n - \frac{f(x_n)}{f'(x_n)},$$

它正好就是第 522 页牛顿法中近似式 (10) 的序列. 关于误差 $x_n - \xi = h_n$，我们有估计

$$|h_{n+1}| = |\phi(x_n) - \phi(\xi)| \leqslant q h_n,$$

其中 q 是以 ξ 和 x_n 为端点的区间内 $|\phi'(x)|$ 的最大值. 因为这里

$$\phi'(x) = \frac{f(x)f''(x)}{f'^2(x)},$$

而 $f(x) = f(x) - f(\xi) = f'(\eta)(x - \xi)$, 我们看到 q 自身与 h_n 同阶,因而就再次证实牛顿法中近似的二次特性.

另一种对 $c(x)$ 的最简单选法是取常数值 $1/f'(x_0)$, 导致递推公式

$$x_{n+1} = \phi(x_n) = x_n - \frac{f(x_n)}{f'(x_0)}.$$

这里, $\phi'(\xi) = 1 - f'(\xi)/f'(x_0)$. 如果 f' 连续而不为 0, 并且我们的初始近似值 x_0 已选得充分接近于解 ξ, 使得

$$|\phi'(\xi)| = \frac{|f'(x_0) - f'(\xi)|}{|f'(x_0)|} < 1,$$

那么我们有一个吸引性不动点 ξ. 这个迭代序列比牛顿法中所用的稍为简单些;然而,收敛将慢得多,有如一个几何级数,象多数迭代模式一样.

例. 作为一个例,我们讨论三次方程

$$f(x) = x^3 - 2x - 5 = 0.$$

因为 $f(2) = -1 < 0$, $f(3) = 16 > 0$, 肯定在区间 $2 < x < 3$ 内存在一个根 ξ. 又因为 $f'(x) = 3x^2 - 2 > 3(2)^2 - 2 > 0$, 所以这区间也只包含一个根. 用牛顿法,从近似值 $x_0 = 2$ 开始, 我们逐次求得

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 2 - \frac{-1}{3(2)^2 - 2} = 2.1, \quad f(x_1) = 0.061,$$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = 2.1 - \frac{0.061}{3(2.1)^2 - 2} = 2.094568.$$

因为 $f(2.1) > 0$, $f(2) < 0$, 根 ξ 位于 2 和 2.1 之间. 在区间 $1.9 < x < 2.2$ 内,我们有估计

$$|f''(x)| = |6x| < 6(2.2) = 13.2,$$

$$f'(x) = 3x^2 - 2 > 3(1.9)^2 - 2 = 8.83.$$

这估计自然对区间 $\xi-0.1<x<\xi+0.1$ 更成立，由此得出[见第 523 页(11)]

$$|\xi-x_{n+1}|\leqslant\frac{13.2}{2(8.83)}|x_n-\xi|^2<0.75|x_n-\xi|^2,$$

只要 $|x_n-\xi|<0.1$. 因为 $|x_0-\xi|=|\xi-2|<0.1$，我们逐次求得

$$|x_1-\xi|<(0.75)(0.1)^2=0.0075,$$
$$|x_2-\xi|<(0.75)(0.0075)^2<0.000042.$$

如果这个近似程度不够，我们取进一步的近似值 x_3，具有误差 $<(0.75)(0.000042)^2<0.0000000013$.

显然，从 f' 和 f'' 都是正的这个事实得知，x_0 之后的所有 x_n 都必定大于 ξ，由此推出

$$h_{n+1}=-f''(\eta)h_n^2/2f'(x_n)<0.$$

对值 x_0，x_1 换用假位法 [第 524 页(13)]，我们求得连接点 $(x_0, f(x_0))$和 $(x_1, f(x_1))$ 的割线与 x 轴的交点 $\bar{\xi}$，

$$\bar{\xi}=x_0-\frac{f(x_0)(x_1-x_0)}{f(x_1)-f(x_0)}=2.09425\cdots.$$

因为曲线在问题所涉及的区间内是下凸的，故割线位于曲线之上，因而近似值 $\bar{\xi}$ 必小于根 ξ.

作为第二个例，让我们解方程

$$f(x)=x\log_{10}x-2=0.$$

我们有 $f(3)=-0.6$，$f(4)=+0.4$，因此，用 $x_0=3.5$ 作为第一个近似值. 使用十位对数表，我们得到逐次近似值:

$$x_0=3.5,\qquad x_1=3.598,$$
$$x_2=3.5972849,\qquad x_3=3.5972850235.$$

附　　录

*A.1 斯特林公式

在许多应用中，特别是在统计学和概率论中，我们发现必需对

$n!$ 有一个象 n 的初等函数一样简单的近似式． 这样一个式子由下述定理给出． 定理是以它的发现者——斯特林(Stirling)命名的(也参看第 667 页第八章)．

当 $n\to\infty$ 时，

$$\frac{n!}{\sqrt{2\pi}\,n^{n+1/2}e^{-n}}\to 1; \tag{14}$$

更确切地说

$$\sqrt{2\pi}\,n^{n+1/2}e^{-n}<n!<\sqrt{2\pi}n^{n+1/2}e^{-n}\left(1+\frac{1}{4n}\right). \tag{14a}$$

换句话说，当 n 的值很大时，表达式 $n!$ 与 $\sqrt{2\pi}\,n^{n+1/2}e^{-n}$ 之差仅仅是一个小的百分数，——如人们所说，两个式子渐近地相等，同时，因子 $1+1/4n$ 给出了近似式的精确度的估计．

如果试图求出曲线 $y=\log x$ 下的面积，我们就会得到这个有名的公式．[1] 用积分法 (第 294 页)，我们求得在坐标为 $x=1$ 和 $x=n$ 之间．这条曲线下的准确面积 A_n 是

$$A_n=\int_1^n \log x dx=x\log -x\Big|_1^n=n\log n-n+1. \tag{15}$$

但是，如果我们用梯形法估计这个面积，在 $x=1$，$x=2$，$\cdots$，$x=n$ 处树立纵坐标，如图 6.8 所示，我们就得到这个面积的近似值 T_n [参看第 508 页(6)]，

$$T_n=\log 2+\log 3+\cdots+\log(n-1)+\frac{1}{2}\log n$$

$$=\log n!-\frac{1}{2}\log n. \tag{16}$$

如果我们作这个合理的假定：A_n 和 T_n 是同一数量级的，我们马上发现 $n!$ 与 $n^{n+1/2}e^{-n}$ 是同一数量级的，本质上这就是斯特林公式所说的．

为了使这个推论确切，我们先证明差 $a_n=A_n-T_n$ 是有界的，由此将直接得出 $T_n=A_n(1-a_n/A_n)$ 与 A_n 是同一数量级

1) 在这里使用的方法是第八章第 661 页将要讨论的欧拉-麦克劳林公式的特例．

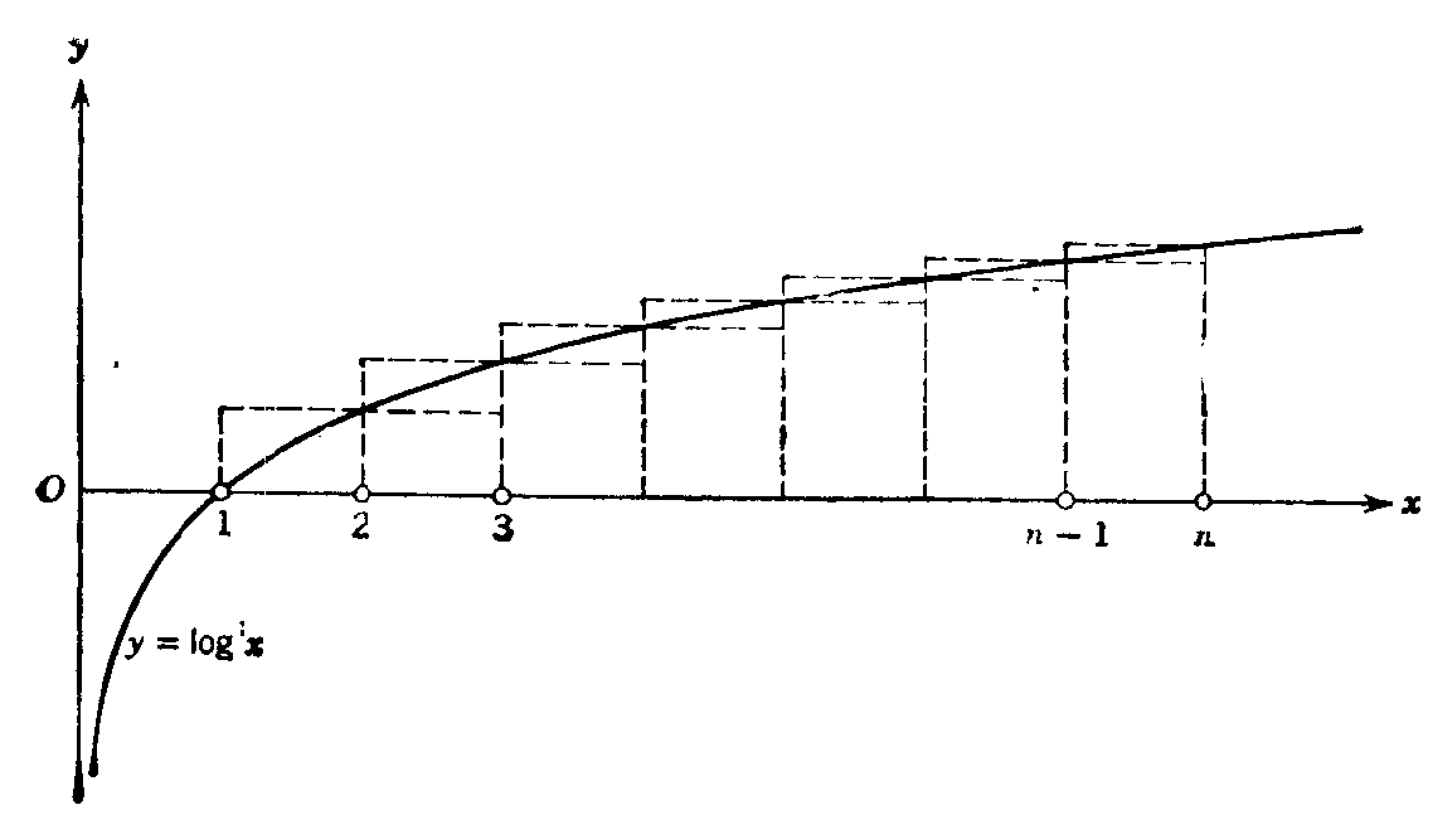


图 6.8

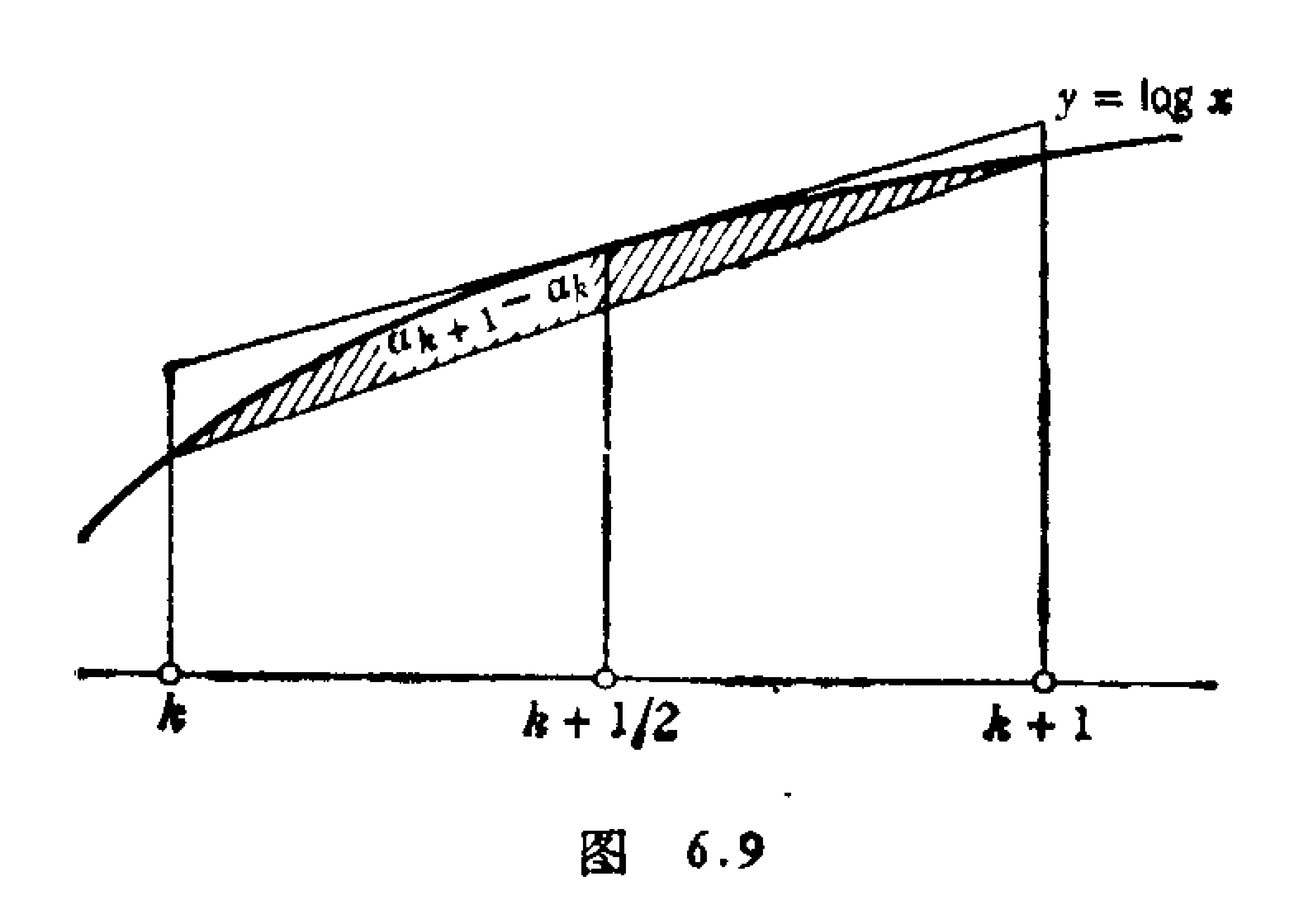


图 6.9

的. 差 $a_{k+1}-a_k$ 是带形 $k<x<k+1$ 中曲线下的面积与割线下的面积之差. 因为曲线是上凸的，位于割线之上，所以 $a_{k+1}-a_k$ 是正的，因而，

$$a_n=(a_n-a_{n-1})+(a_{n-1}-a_{n-2})+\cdots+(a_2-a_1)+a_1$$

单调增加. 另外，差 $a_{k+1}-a_k$ 显然小于在 $x=k+1/2$ 点切线下的面积与割线下的面积之差(参看图 6.9)；因此，我们有不等式

$$a_{k+1}-a_k<\log\left(k+\frac{1}{2}\right)-\frac{1}{2}\log k-\frac{1}{2}\log(k+1)$$

$$=\frac{1}{2}\log\left(1+\frac{1}{2k}\right)-\frac{1}{2}\log\left[1+\frac{1}{2(k+1/2)}\right]$$

$$< \frac{1}{2}\log\left(1+\frac{1}{2k}\right)-\frac{1}{2}\log\left[1+\frac{1}{2(k+1)}\right].$$

把这些不等式对 $k=1, 2, \cdots, n-1$ 加起来，我们发现右边除两项外所有项都互相抵消，因而（因 $a_1=0$）有

$$a_n<\frac{1}{2}\log\frac{3}{2}-\frac{1}{2}\log\left(1+\frac{1}{2n}\right)<\frac{1}{2}\log\frac{3}{2}.$$

因为 a_n 有界，而且单调递增，当 $n\to\infty$ 时，趋向一个极限 a. 关于 $a_{k+1}-a_k$ 的不等式，现在给出

$$a-a_n=\sum_{k=n}^{\infty}(a_{k+1}-a_k)<\frac{1}{2}\log\left(1+\frac{1}{2n}\right).$$

因为由定义 $A_n-T_n=a_n$，从 (15)，(16)，我们有

$$\log n!=1-a_n+\left(n+\frac{1}{2}\right)\log n-n,$$

或记 $\alpha_n=e^{1-a_n}$，

$$n!=\alpha_n n^{n+1/2}e^{-n}.$$

序列 α_n 是单调递减的，并且趋向于极限 $\alpha=e^{1-a}$；因而

$$1<\frac{\alpha_n}{\alpha}=e^{a-a_n}<e^{(1/2)\log(1+1/2n)}$$

$$=\sqrt{1+\frac{1}{2n}}<1+\frac{1}{4n}.$$

于是我们有

$$\alpha n^{n+1/2}e^{-n}<n!<\alpha n^{n+1/2}e^{-n}\left(1+\frac{1}{4n}\right),$$

剩下的事情只是寻找极限 a 的确实值. 我们用第三章第 300 页公式(80)：

$$\sqrt{\pi}=\lim_{n\to\infty}\frac{(n!)^2 2^{2n}}{(2n)!\sqrt{n}}.$$

用 $\alpha_n n^{n+1/2}e^{-n}$ 代替 $n!$，用 $\alpha_{2n}2^{2n+1/2}n^{2n+1/2}e^{-2n}$ 代替 $(2n)!$，我们立即得到

$$\sqrt{\pi}=\lim_{n\to\infty}\frac{\alpha_n^2}{\alpha_{2n}\sqrt{2}}=\frac{\alpha^2}{\alpha\sqrt{2}},$$

因此，$\alpha = \sqrt{2\pi}$. 这样，斯特林公式证明完毕.

斯特林公式不仅有它的理论价值，它还是对大数 n 作 $n!$ 的数值计算的一个很有用的工具. 我们不用再去计算数目极大的整数连乘，而只要借助于对数用斯特林公式作很少的运算就可以了. 譬如 $n = 10$，我们得到斯特林表达式的值 3598696（用 7 位数表），而 10! 的准确值则是 3628800，百分误差仅仅是 5/6%.

问　　题

6.1 节，第 506 页

1．证明当 $f''(x) \geqslant 0$ 时，梯形法给出一个比 f 的准确积分较大的值；而切线法给出一个较小的值.

2．设要用辛卜生法则计算下列积分，要求准确到 p 位小数，试估计所需 $h=(b-a)/n$ 的值：

$$(a)\ \log 2 = \int_1^2 \frac{1}{x}\,dx, \qquad (b)\ \pi = 4\int_0^1 \frac{1}{1+x^2}\,dx.$$

3．用 k 和 s（$k<1$，$s<1$）表示，估计计算椭圆积分

$$u(s) = \int_0^s \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}$$

时误差不超过 ϵ 所需的点数.

4．设 $f(x)$ 在区间 $\alpha \leqslant x \leqslant \alpha + h$ 上连续，并且导数一致有界，即存在常数 M_1，使得 $|f'(x)| \leqslant M_1$. 证明对任一固定点 ξ，$\alpha \leqslant \xi \leqslant \alpha + h$，都有估计式

$$\left| \int_\alpha^{\alpha+h} f(x)dx - hf(\xi) \right| \leqslant \frac{M_1 h^2}{2}.$$

5．计算 $\int_a^\infty e^{-x^2}dx$ 的数值，使误差不超过 1/100.

6.2 节，第 515 页

1．一个摆的周期如下：

$$T = 2\pi\sqrt{\frac{l}{g}},$$

其中 l 是摆长. 如果用此摆驱动一个钟表，钟表每天快一分钟，试确定 l 所必需的修正值.

2．为测量一山的高，从地面观察山顶上的一个一百米高的塔. 塔脚的

仰角为 42°，塔本身张角为 6°. 在这 42° 角的误差为 1° 的条件下，问所确定的山高的误差限度是多少?

6.3 节，第 520 页

1. (a) 为要解方程 $x = f(x)$，试确定如何选常数 a 最好，使得迭代模式

$$x_{k+1} = x_k + a[x_k - f(x_k)]$$

收敛速度在解的邻域内尽可能地快.

(b) 用这个方法求解关于 $\sqrt{A}$ 的方程，

$$x = \frac{A}{x}.$$

(c) 试证当 $A \geqslant 1$ 时，在(b)中得到的迭代模式的每一步准确的小数位数至少是成倍增加的.

2. (a) 试确定多项式

$$g(x) = a + bx^2$$

的最好选取，使得对于 $\sqrt{A}$ 的迭代模式

$$x_{k+1} = x_k + g(x_k)\left(x_k - \frac{A}{x_k}\right)$$

在解的邻域内收敛最快.

(b) 估计收敛速度.

(c) 如何适当选择更高次的多项式 $g(x)$，进一步改进收敛性.

3. 研究问题 1 和 2 类型的适当模式用于计算 $\sqrt[n]{A}$

A.1 节，第 530 页

1. 证明 $\lim\limits_{n\to\infty} \frac{\sqrt[n]{n!}}{n} = \frac{1}{e}$.

2* 通过考察 $\int_{1/2}^{n+1/2} \log(\alpha + x)dx$, $\alpha > 0$, 试证明

$$\alpha(\alpha + 1)\cdots(\alpha + n) = a_n n! n^{\alpha},$$

其中 a_n 有一个大于 0 的下界. 进一步证明，对 n 充分大的值，a_n 单调递减. [当 $n \to \infty$ 时，a_n 的极限是 $1/\Gamma(\alpha)$]

3. 试求 $\log \frac{n_1! n_2! \cdots n_l!}{n!}$ 的一个近似表达式，其中 $n_1 + n_2 + \cdots + n_l = n$.

4. 证明 $\frac{1}{\sqrt{1-x}}$ 的二项展式中 x^n 的系数由 $\frac{1}{\sqrt{\pi n}}$ 渐近地给出.

第七章　无穷和与无穷乘积

几何级数、泰勒级数以及本书前面讨论过的一些例都启示我们可以从更一般的观点出发来好好地研究分析中的那些极限过程，其中包括无穷级数的求和．原则上说，任何极限值

$$S = \lim_{n \to \infty} s_n$$

都能够写成一个无穷级数；我们只需对 $n > 1$ 设 $a_n = s_n - s_{n-1}$，并设 $a_1 = s_1$，就得到

$$s_n = a_1 + a_2 + \cdots + a_n,$$

于是，当 n 增加时数值 S 就作为 n 项和 s_n 的极限而出现．我们称 S 是"无穷级数

$$a_1 + a_2 + a_3 + \cdots$$

的和"来表示这一事实．

这样一个"无穷和"只是表示极限的一种方法，在这个极限中，每一个逐次的近似值是由前一个近似值再加一项得到的．例如，一个数的十进小数表达式在原则上就仅仅是把这个数 a 表示成无穷级数 $a = a_1 + a_2 + a_3 + \cdots$ 的形式，在这里，如果 $0 \leqslant a \leqslant 1$，那末项 a_n 就可用 $\alpha_n \times 10^{-n}$ 来替换，而 α_n 是一个 0 到 9 之间的整数．

由于每一个极限值都能够写成一个无穷级数的形式，级数的专门研究似乎是多余的．然而，经常出现这种情况，极限值天然地以这样的无穷级数的形式出现，而这种无穷级数呈现出特别简单的形成规律．并不是每一个级数都有一个容易认识的形成规律的．例如，数 π 无疑能表示成一个十进小数(这个小数是一个级数 $\sum c_\nu 10^{-\nu}$)；但是我们知道，没有一个简单的规律使我们能够说出这个小数的任意一位的值，譬如说第 7000 位的值．然而，如果我

们考虑用 $\frac{\pi}{4}$ 的莱布尼兹-格雷戈里级数来代替，我们就有一个具有完全清楚的一般的形成规律的表达式 [参看第 468 页 (7)].

类似于无穷级数的是无穷乘积，在无穷级数里极限的近似值是由反复加新的项形成的；在无穷乘积里极限的近似值是由反复乘新的因子形成的．然而，我们将不深入到无穷乘积的一般理论中，这一章和第八章的主要课题将是无穷级数．

7.1 收敛与发散的概念

a．基本概念

柯西收敛准则．我们考虑具有"一般项" a_n 的一个无穷级数，于是这个级数[1]就具有形式

$$a_1 + a_2 + \cdots = \sum_{\nu=1}^{\infty} a_\nu.$$

右边带有求和号的记号只是左边表达式的一种缩写方式．

如果当 n 增加时，第 n 个部分和

$$s_n = a_1 + a_2 + \cdots + a_n = \sum_{\nu=1}^{n} a_\nu$$

趋向于一个极限

$$S = \lim_{n\to\infty} s_n$$

我们就说这个级数是收敛的；否则我们就说它是发散的．在收敛的情况下，我们称 S 为级数的和．

我们早已遇到过许多收敛级数的例子；例如，几何级数 $1 + q + q^2 + \cdots$，当 $|q| < 1$ 时，收敛到和 $1/(1-q)$，还有 $\log 2$ 的级数，e 的级数等等．

用无穷级数的语言，柯西收敛判别法(参看第一章第 78 页)表示如下：

---

1) 由于形式上的理由，我们包括了某些数 a_n 可以是零这种可能性．如果从某一个指标数 N 以后(即当 $n > N$ 时)所有的项都为零，我们称此级数为有尽级数．

一个级数 $\sum_{\nu=1}^{\infty} a_\nu$ 收敛的必要充分条件是：如果 m 和 n 取得充分大时，数

$$|s_m - s_n| = |a_{n+1} + a_{n+2} + \cdots + a_m| \tag{1}$$

$(m > n)$ 就变得任意小．换句话说：当且仅当下面的条件被满足时，级数是收敛的：对于一个给定的正数 ε，可以找到一个指标数 $N = N(\varepsilon)$，使得只要 $m > N$ 和 $n > N$，上面的表达式 $|s_m - s_n|$ 就小于 ε.

我们能够用 $q = \frac{1}{2}$ 的几何级数来说明这个收敛判别法． 如果我们选择 $\varepsilon = \frac{1}{10}$，我们只需取 $N = 4$．因为

$$|s_m - s_n| = \frac{1}{2^n} + \cdots + \frac{1}{2^{m-1}}$$

$$= \frac{1}{2^{n-1}}\left(\frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^{m-n}}\right) < \frac{1}{2^{n-1}},$$

所以当 $n > 4$ 时

$$\frac{1}{2^{n-1}} < \frac{1}{10}.$$

如果我们选择 ε 等于 $\frac{1}{100}$，取 7 作为 N 的对应值就足够了，这是容易验证的．

显然，一个级数收敛的必要条件是

$$\lim_{n\to\infty} a_n = 0.$$

否则，对于 $m = n + 1$ 收敛准则就无疑不能满足．但是这个必要条件对于收敛而言决不是充分的；相反，容易找到一个无穷级数，它的一般项 a_n 当 n 增加时趋向于 0，但是它的和不存在，因为当 n 增加时，部分和 s_n 递增但没有极限．

例．一个例是级数

$$1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \cdots + \frac{1}{\sqrt{n}} + \cdots,$$

它的一般项为 $\frac{1}{\sqrt{n}}$．我们立即看到

$$s_n > \frac{1}{\sqrt{n}} + \cdots + \frac{1}{\sqrt{n}} = \frac{n}{\sqrt{n}} = \sqrt{n}.$$

当 n 增加时，第 n 个部分和递增并趋向无穷大，因此级数发散.

对于调和级数

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots$$

这一经典例子，同样有上述结论. 这里

$$a_{n+1} + \cdots + a_{2n} = \frac{1}{n+1} + \cdots + \frac{1}{2n}$$

$$> \frac{1}{2n} + \cdots + \frac{1}{2n} = \frac{1}{2}.$$

由于 n 和 $m = 2n$ 可以要多大就选得多大，柯西判别法得不到满足，所以级数发散；事实上，因为所有的项都是正的，其第 n 个部分和明显地趋于无穷. 另一方面，由同样的数但采取交错的符号，所形成的级数

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - + \cdots + \frac{(-1)^{n-1}}{n} + \cdots$$

收敛（参看第五章第 466 页 (4)），并有和 log 2.

不能认为每个发散级数 s_n 都趋于 $+\infty$ 或 $-\infty$. 例如在级数

$$1 - 1 + 1 - 1 + 1 + - \cdots$$

中，我们看到它的部分和 s_n 交替地取值 1 和 0，并且由于这种前后的摆动，既不趋向一个确定的极限，又不在数值上无限增加.

下面的事实，虽然是不言自明的，但是非常重要，应该引起注意. 一个级数的收敛性或发散性，决不因加进或去掉有限项而变更. 就收敛性或发散性而论，不论从 a_0 或 a_1 或 a_5 或任意选择的任一其它项开始作为级数的首项，都是毫无关系的.

b. 绝对收敛与条件收敛

调和级数 $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots$ 是发散的，但是如果我们

每隔一项改变其符号，则所成的 log 2 的级数收敛．另一方面，几何级数 $1-q+q^2-q^3+-\cdots$ 在 $0\leqslant q<1$ 的条件下收敛且有和 $1/(1+q)$，并且使所有的符号为正，我们得到级数

$$1+q+q^2+q^3+\cdots,$$

这个级数也是收敛的，它有和 $1/(1-q)$.

这里呈现一种差别，我们必须考察．所有的项都为正的级数只有两种可能情况，或者收敛或者当 n 增加时部分和递增无界．因为部分和是一个单调递增序列，如果部分和有界，必然收敛．当 n 增加时，如果一般项趋于零的速度足够快，级数就收敛；另一方面，如果一般项根本不趋向于零或趋向于零的速度太慢，级数就发散．然而在级数既有正项又有负项时，符号的改变正是造成收敛的原因，当正项的部分和增加太大时，负项进行补偿，致使最终结果趋向一个确定的极限．

为了有可能更好地理解，我们考察既有正项又有负项的级数 $\sum_{\nu=1}^{\infty} a_\nu$，将它与另一个与它有相同的项，符号全为正的级数，即

$$|a_1|+|a_2|+\cdots=\sum_{\nu=1}^{\infty}|a_\nu|,$$

进行比较，如果后面这个级数收敛，于是当 n 充分大，且 $m>n$ 时，表达式

$$|a_{n+1}|+|a_{n+2}|+\cdots+|a_m|$$

无疑将任意小；由于关系式

$$|a_{n+1}+\cdots+a_m|\leqslant|a_{n+1}|+\cdots+|a_m|,$$

左边的表达式也是任意小的，因此根据哥西判别法，原先的级数 $\sum_{\nu=1}^{\infty} a_\nu$ 收敛．在这种情况下，原先的级数称为**绝对收敛**．它的收敛性是由于它的项绝对地小，而与它的符号的变化无关．

另一方面，如果具有一般项 $|a_n|$ 的级数发散，而原先的级数仍然收敛，我们称原先的级数为**条件收敛**．条件收敛起因于具有相反符号的项互相补偿．

莱布尼兹判别法. 对于条件收敛的级数，莱布尼兹收敛判别法是经常有用的：

假如一个级数的项的符号正负交错，而且项的绝对值 $|a_n|$ 单调趋于零(因此 $|a_{n+1}| \leqslant |a_n|$)，则级数 $\sum_{\nu=1}^{\infty} a_\nu$ 收敛. [例如：莱布尼兹级数，见第 468 页(7)].

为了证明这个判别法，不失一般性，我们设 $a_1 > 0$，并将此级数写成

$$b_1 - b_2 + b_3 - + \cdots.$$

现在式中所有的项 b_n 均为正数，b_n 趋于零，并且满足条件 $b_{n+1} \leqslant b_n$. 如果我们用两种不同的方法组合项，

$$b_1 - (b_2 - b_3) - (b_4 - b_5) - \cdots$$

及

$$(b_1 - b_2) + (b_3 - b_4) + (b_5 - b_6) + \cdots,$$

我们立即看到部分和 $s_n = \sum_{\nu=1}^{n} a_\nu$ 满足下面两个关系式

$$s_1 \geqslant s_3 \geqslant s_5 \geqslant \cdots \geqslant s_{2n+1} \geqslant \cdots,$$

$$s_2 \leqslant s_4 \leqslant s_6 \leqslant \cdots \leqslant s_{2n} \leqslant \cdots.$$

另一方面，$s_{2n} \leqslant s_{2n+1} \leqslant s_1$ 并且 $s_{2n+1} \geqslant s_{2n} \geqslant s_2$. 所以奇数的部分和 $s_1, s_3, \cdots$ 形成一个单调递降的序列，它以 s_2 为下界；因

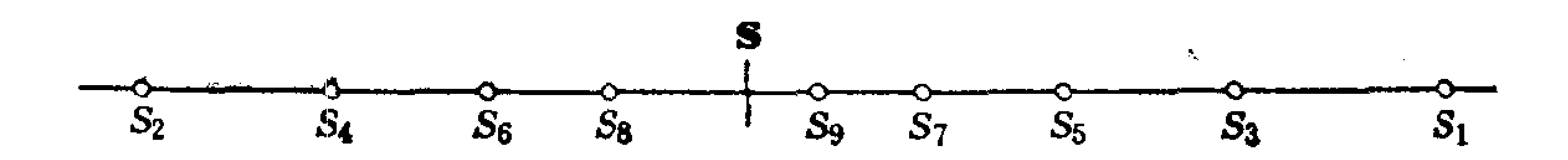


图 7.1 一个交错级数的收敛性

此这个序列有一个极限 L (第77页). 偶数的部分和 $s_2, s_4, \cdots$ 类似地形成一个单调递增的序列，它以 s_1 为上界，因此这个序列必有一极限值 L'. 由于数 s_{2n} 与 s_{2n+1} 互相只相差 b_{2n+1}，当 n 增加时 b_{2n+1} 趋于 0，于是极限值 L 和 L' 彼此相等. 即，偶的和奇的部分和趋于同一个极限，这个极限我们现在记作 S (参看图

7.1). 然而，这蕴含着我们的级数是收敛的，它的和是 S. 这正是要证的结论.

*阿贝尔判别法

阿贝尔收敛判别法是一个条件收敛级数的判别法,它包括莱布尼兹判别法作为一个特殊情形. 设 $a_1 + a_2 + \cdots$ 是一个无穷级数,它的部分和 $s_n = a_1 + \cdots + a_n$ 有不依赖于 n 的界.又设 $p_1, p_2, \cdots$ 是一个单调递降到零的正数序列. 于是无穷级数

$$p_1a_1 + p_2a_2 + \cdots \qquad (2)$$

收敛.(对于特殊的级数 $a_1 + a_2 + \cdots = +1 - 1 + 1 - 1 + - \cdots$,我们得到 $p_1 - p_2 + p_3 - \cdots$收敛,这就是莱布尼兹判别法.)证明是不难得到的,如果我们应用哥西判别法,借助"分部求和法"我们有估计

$$\begin{aligned}
&|p_{n+1}a_{n+1} + p_{n+2}a_{n+2} + \cdots + p_m a_m| \\
&= |p_{n+1}(s_{n+1} - s_n) + p_{n+2}(s_{n+2} - s_{n+1}) + \cdots + p_m(s_m - s_{m-1})| \\
&= |-p_{n+1}s_n + p_m s_m + (p_{n+1} - p_{n+2})s_{n+1} + (p_{n+2} - p_{n+3})s_{n+2} + \cdots \\
&\quad + (p_{m-1} - p_m)s_{m-1}| \\
&\leqslant p_{n+1}M + p_m M + (p_{n+1} - p_{n+2} + p_{n+2} - p_{n+3} + \cdots + p_{m-1} - p_m)M \\
&= 2p_{n+1}M,
\end{aligned}$$

其中 M 是 $|s_i|$ 的一个界;因为 $p_{n+1} \to 0$, 于是根据哥西判别法就可推出级数(2)的收敛性.

*最后,关于绝对收敛级数和条件收敛级数之间的基本差别,我们作另一个一般的评论. 我们考虑一个收敛级数 $\sum_{\nu=1}^{\infty} a_\nu$. 我们用 $p_1, p_2, p_3, \cdots$ 记级数的正项,同时用 $-q_1, -q_2, -q_3, \cdots$ 记级数的负项. 假如我们组成给定级数的第 n 个部分和 $s_n = \sum_{\nu=1}^{n} a_\nu$, 其中必然出现某个数量的正项,譬如说是 n', 某个数量的负项,譬如说是 n'',这里 $n' + n'' = n$. 而且,如果级数中正项的数目和负项的数目一样是无限多,那末两数 n' 和 n'' 将像 n 一样递增无界. 我们立即看到,部分和 s_n 恰恰等于级数的正项部分和 $\sum_{\nu=1}^{n'} p_\nu$ 加上负项的部分和 $-\sum_{\nu=1}^{n''} q_\nu$. 如果给定的级数绝对收敛,那末正项的级数 $\sum_{\nu=1}^{\infty} p_\nu$ 和负项的绝对值的级数 $\sum_{\nu=1}^{\infty} q_\nu$ 必定同时收敛. 因为

当 m 增加时，部分和 $\sum_{\nu=1}^{m} p_\nu$ 和 $\sum_{\nu=1}^{m} q_\nu$ 是单调非减的序列，有上界 $\sum_{\nu=1}^{\infty} |a_\nu|$.

于是一个绝对收敛的级数的和恰好等于只由正项组成的级数的和加上只由负项组成的级数的和，或者，换句话说，是等于两个具有正项的级数之差.

因为 $\sum_{\nu=1}^{n} a_\nu = \sum_{\nu=1}^{n'} p_\nu - \sum_{\nu=1}^{n''} q_\nu$; 当 n 增加时，n' 和 n'' 也递增无界，因此左边的极限必须等于右边两个和的差. 如果级数只包含一种特定符号的有限项，情况是相当简单的. 另一方面，如果级数不是绝对收敛，而是条件收敛，则级数 $\sum_{\nu=1}^{\infty} p_\nu$ 和 $\sum_{\nu=1}^{\infty} q_\nu$ 必须同时发散. 因为假如它们两者都收敛，级数将绝对收敛，与我们的假设矛盾. 假如只是一个发散，譬如是 $\sum_{\nu=1}^{\infty} p_\nu$，而另一个收敛，于是分离成正和负两个部分，$s_n = \sum_{\nu=1}^{n'} p_\nu - \sum_{\nu=1}^{n''} q_\nu$，这表明这个级数不可能收敛；因为当 n 递增时，n' 和 $\sum_{\nu=1}^{n'} p_\nu$ 将递增无界，而项 $\sum_{\nu=1}^{n''} q_\nu$ 将趋向一个有穷的极限，因此部分和 s_n 将递增无界.

因此，我们看到，不能把一个条件收敛的级数看作两个收敛级数之差，其中一个由它的正项组成，同时另一个由它的负项的绝对值组成.

与这个事实紧密相联的是绝对收敛和条件收敛之间的另一个差别，我们现在就简要地提一提.

*c. 项的重新排列

有穷和的一个性质，即我们能够改变项的次序，或者，如我们所说的，任意排列项的次序不改变和的值. 问题在于：在一个无穷级数里什么是项的次序的改变的确切意义？并且作这样一个重新排列能使和的值不改变吗？虽然在有穷和里做到这些没有困难，例如，按相反的次序加这些项，而在无穷

级数里这样一种可能性就不存在，它没有最后一项可以用它去开头．现在在一个无穷级数里次序的改变只能是这样的意思：我们说一个级数 $a_1+a_2+a_3+\cdots$ 经重新排列而变换为一个级数 $b_1+b_2+b_3+\cdots$，只要第一个级数里的每一项 a_n 确实在第二个级数里出现一次，并且反过来也是这样．举例来说，被代替的 a_n 的总数可以像 n 一样无限增大，但唯一的一点是 a_n 必须出现在新的级数里的某个地方．如果有些项被移到级数的较后的位置，当然，另一些项必须被移到较前的位置．例如，级数

$$1+q+q^2+q^4+q^3+q^8+q^7+q^6+q^5+q^{16}+\cdots$$

是几何级数 $1+q+q^2+\cdots$ 的一个重新排列[1]．

关于次序的改变绝对收敛级数和条件收敛级数之间存在一个基本的差别．

在绝对收敛级数里项的重新排列不影响其收敛性，并且级数和的值不改变，与有限和一样地正确．

另一方面，在条件收敛级数里，级数的适当的重新排列就能够任意改变级数和的值，并且，如果需要，甚至能够使级数发散．

首先，关于绝对收敛级数是容易确立的．开始我们假设，级数只有正项，并且考虑第 n 个部分和 $s_n=\sum_{\nu=1}^{n}a_\nu$．只要 m 选得足够大，这个部分和的所有项出现在重排级数的第 m 个部分和 $t_m=\sum_{\nu=1}^{m}b_\nu$ 中，因此 $t_m\geqslant s_n$．另一方面，我们能够决定一个足够大的指标数 n' 使得第一个级数的部分和 $s_{n'}=\sum_{\nu=1}^{n'}a_\nu$ 包含项 $b_1,b_2,\cdots,b_m$ 的全体．于是由此得出 $t_m\leqslant s_{n'}\leqslant A$，其中 A 是第一个级数的和．因此对于 m 的所有充分大的值，我们有 $s_n\leqslant t_m\leqslant A$，同时因为 s_n 与 A 的差能够是任意小的数，由此可见重排级数也是收敛的，事实上，与原来的级数一样趋向同一个极限 A．

如果绝对收敛级数既有正项又有负项，可以把它看作两个只有正项的级数的差．由于在原来的级数的重新排列的过程中，这两个级数的每一个只是项交换了位置，因此与未交换位置的级数一样收敛到同样的值，于是原来的级数经过重新排列以后，级数和的值同样不改变．根据刚才讨论的情况，这新的级数是绝对收敛的，因此它是两个正项重排级数的差．

对于初学者来说，刚才证明的事实似乎是显然的．不具备这种性质的条

1) 对于每一个 $n>0$，相应于 $\alpha^n<k\leqslant\alpha^{n+1}$ 的这些项 q^k，按相反的次序写．

件收敛级数的一个例能够说明这个事实的确是需要证明的，并且在证明过程中绝对收敛是本质的。我们取熟悉的 $\log 2$ 的级数，把它用因子 $\frac{1}{2}$ 乘的结果写在它的下面，

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\frac{1}{6}+\frac{1}{7}-\frac{1}{8}+-\cdots=\log 2,$$

$$\frac{1}{2}\qquad -\frac{1}{4}\qquad +\frac{1}{6}\qquad -\frac{1}{8}+-\cdots=\frac{1}{2}\log 2,$$

并且把它们加起来，按纵列结合项[1]. 于是我们得到

$$1+\frac{1}{3}-\frac{1}{2}+\frac{1}{5}+\frac{1}{7}-\frac{1}{4}+\frac{1}{9}+\frac{1}{11}-\frac{1}{6}++-\cdots=\frac{3}{2}\log 2.$$

后面这一级数明显地能够由原来的级数经过重排而得到，但是级数的和的值已经被乘上了因子 $\frac{3}{2}$. 不难想像，这个表面上看起来似是而非的发现，对于习惯于用无穷级数进行运算，而不考虑它们的收敛性的十八世纪的数学家产生了怎样的影响.

*我们将给出以上关于条件收敛级数 $\sum a_n$ 由于项的次序的改变而引起和的改变的定理的证明，虽然我们没有机会去利用这个结果. 设 $p_1, p_2, \cdots$ 是级数的正项，而 $-q_1, -q_2, \cdots$ 是级数的负项. 由于当 n 增加时，绝对值 $|a_n|$ 趋于 0，数 p_n 和 q_n 当 n 增加时也必须趋于 0. 然而，正像我们早已看到的，和 $\sum_{\nu=1}^{\infty} p_\nu$ 一定发散，并且 $\sum_{\nu=1}^{\infty} q_\nu$ 也同样发散.

现在我们能够容易地找到原来级数的一个重新排列，重排后的级数以任意的数 a 作为它的和. 为了明确起见，假设 a 是正的. 于是我们把前面 n_1 个正项加在一起，使得和 $\sum_{\nu=1}^{n_1} p_\nu$ 刚好大于 a. 由于和 $\sum_{1}^{n_1} p_\nu$ 随着 n_1 递增无界，因此取足够多的项使得部分和大于 a 总是可能的. 于是部分和与精确值 a 的差至多是 p_{n_1}. 现在我们加足够多的负项 $-\sum_{1}^{m_1} q_\nu$，使得和 $\sum_{1}^{n_1} p_\nu-\sum_{1}^{m_1} q_\nu$ 刚好小于 a；根据级数 $\sum_{1}^{\infty} q_\nu$ 的发散性，这也是可能的. 这时这个和与 a 之

1) 关于级数的加法参看 7.1d 节.

间的差至多是 q_{m_1}. 现在我们加另外的足够多的正项 $\sum_{n_1+1}^{n_2} p_\nu$，使得部分和再一次刚好大于 a，由于该正项级数发散，这还是可能的. 这时部分和与 a 之间的差至多是 p_{n_2}. 我们再一次加足够多的负项 $-\sum_{m_1+1}^{m_2} q_\nu$，这里开始的项是在前面用过的最后一项的下一项，使得和再一次刚好小于 a，并且反复使用同样的方法. 这样得到的和的值将围绕数 a 而摆动，同时当过程持续足够长以后，这个摆动将只在任意狭小的范围之内发生；由于当 ν 足够大时，项 p_ν 和 q_ν 自身趋于 0，于是摆动的区间的长度也将趋于 0，于是定理得证.

同样，我们能够用这样的方法来重排级数使它发散：我们只需把正项的数目取得如此之大，以至它与负项匹配时，补偿不再发生.

d. 无穷级数的运算

显然，两个收敛的无穷级数 $a_1+a_2+\cdots=S$ 和 $b_1+b_2+\cdots=T$ 可以逐项相加，即，由项 $c_n=a_n+b_n$ 形成的级数收敛，并以值 $S+T$ 作为它的和[1]. 由于

$$\sum_{\nu=1}^{n} c_\nu=\sum_{\nu=1}^{n} a_\nu+\sum_{\nu=1}^{n} b_\nu \to S+T.$$

同样显然，如果我们用相同的因子去乘收敛的无穷级数的每一项，这个级数仍然收敛，它的和被乘以这同一个因子.

对于这些运算，收敛是绝对的还是条件的，都是无所谓的. 另一方面，进一步的研究表明，除非两个级数中至少有一个是绝对收敛的；否则，两个无穷级数用有穷和的方法去乘，不一定导至一个收敛级数作为它们的乘积的值(参看附录第 582 页).

7.2 绝对收敛和发散的判别法

在 7.1**b** 节我们已经遇见过对于级数的条件收敛很有用的莱布尼兹判别法. 下面我们将只考虑涉及绝对收敛的准则.

---

1) 这个定理确实无非是两项和的极限等于它们的极限的和这个事实(参看第一章第 75 页)的另一种陈述.

a. 比较判别法. 控制级数

收敛性的所有这种考虑都依赖于所要考虑的级数与第二个级数的比较；这第二个级数是用这样一种方法来选取的，使得它的收敛性能够容易被判别出来. 一般的比较判别法可以叙述如下：

如果数 $b_1, b_2, \cdots$ 都是正的，同时级数 $\sum_{\nu=1}^{\infty} b_\nu$ 收敛，并且如果对于所有的 n 都有

$$|a_n| \leqslant b_n,$$

则级数 $\sum_{n=1}^{\infty} a_n$ 是绝对收敛的.

根据柯西判别法，这个证明几乎是显然的. 因为如果 $m \geqslant n$，我们有

$$|a_n + \cdots + a_m| \leqslant |a_n| + \cdots + |a_m| \leqslant b_n + \cdots + b_m.$$

由于级数 $\sum_{n=1}^{\infty} b_n$ 收敛，只要 n 和 m 充分大，右边就任意小. 由此可见，对于这样的 n 和 m 的值左边也是任意小，因此根据哥西判别法，给定的级数收敛. 这个收敛性是绝对的，因为我们的讨论可以完全同样地应用到绝对值 $|a_n|$ 的级数的收敛性上去.

下述事实的类似证明可以留给读者. 如果

$$|a_n| \geqslant b_n > 0,$$

并且级数 $\sum_{n=1}^{\infty} b_n$ 发散，则级数 $\sum_{n=1}^{\infty} a_n$ 一定不绝对收敛.

有时把上述具有正项 b_n 的级数分别称为相应于具有项 a_n 的级数的上控制级数和下控制级数.

b. 与几何级数相比较的收敛判别法

在比较判别法的应用中，最常用来作为上控制级数的比较级数是几何级数. 于是我们立即得到下面的定理.

定理 级数 $\sum_{n=1}^{\infty} a_n$ 绝对收敛的一个条件是，从某一项起关系

式

$$|a_n| < cq^n \tag{3}$$

成立. 其中 c 是不依赖于 n 的一个正数, q 是任意一个小于 1 的固定的正数.

比值判别法和根式判别法. 上述判别法经常表示成下面较弱的形式之一: 级数 $\sum_{n=1}^{\infty} a_n$ 绝对收敛的一个条件是,从某一项起关系式

$$\left|\frac{a_{n+1}}{a_n}\right| < q \tag{4a}$$

成立, 其中 q 仍是一个小于 1 的正数,并且不依赖于 n, 或者从某一项起关系式

$$\sqrt[n]{|a_n|} < q \tag{4b}$$

成立, 其中 q 是一个小于 1 的正数. 特别地,当关系式

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = k < 1 \tag{5a}$$

或

$$\lim_{n\to\infty}\sqrt[n]{|a_n|} = k < 1 \tag{5b}$$

成立时,就满足上述判别法的条件.

上述断言容易用下述方法建立.

让我们设从下标 n_0 以后,即当 $n > n_0$ 时,比值判别法的准则 (4a) 是满足的. 为简洁起见,我们令 $a_{n_0+m+1} = b_m$, 因而得到

$$|b_1| < q|b_0|, \qquad |b_2| < q|b_1| < q^2|b_0|,$$
$$|b_3| < q|b_2| < q^3|b_0|,$$

等等;因此

$$|b_m| < q^m|b_0|,$$

于是当 $n > n_0$ 并且 $c = q^{-n_0-1}|b_0|$ 时,

$$|a_n| = |b_{n-n_0-1}| < q^{n-n_0-1}|b_0| = cq^n,$$

这就确立了我们的断言. 对于根式判别法的准则 (4b), 我们立刻有 $|a_n| < q^n$, 因而我们的断言随之立即得到.

最后，为了证明准则(5)，我们考虑任意一个数 q，使得 $k < q < 1$. 于是从某一个 n_0 以后，即 $n > n_0$ 时，关系式 5a，b 分别隐含 $\left|\frac{a_{n+1}}{a_n}\right| < q$ 和 $\sqrt[n]{|a_n|} < q$，因为从某项以后 $\left|\frac{a_{n+1}}{a_n}\right|$ 或 $\sqrt[n]{|a_n|}$ 与 k 的差小于 $(q-k)$，于是，基于已经证明的结果，断言成立.

我们强调指出，从初始的准则 $|a_n| < cq^n$ 得到的四个判别法 4a，b，5a，b 并不互相等价也不同初始的准则等价，就是说，他们不能互相推得. 我们很快将从一些例中看到，如果一个级数满足条件中的一个，它不一定满足其它任何一个.

为完满起见，可以指出，如果从某一项起，对于某个正数 c

$$|a_n| > c,$$

或者从某一项起

$$\sqrt[n]{|a_n|} > 1,$$

或者 $$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = k, \quad 或 \quad \lim_{n\to\infty}\sqrt[n]{|a_n|} = k,$$

其中 k 是大于 1 的数，则级数 $\sum_{n=1}^{\infty} a_n$ 一定发散. 因为我们一看就可认出，在这种级数里其项不可能随 n 的增大而趋于零，因此这些级数必定发散.（在这种情况下，级数甚至不可能条件收敛.）

我们的判别法给出了一个级数绝对收敛的充分条件；亦即，当条件满足时，能够得出结论，级数是绝对收敛的. 然而，它们的确不是必要条件；亦即，能够作出不满足这些条件的绝对收敛级数.

譬如，条件

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = 1 \quad 或 \quad \lim_{n\to\infty}\sqrt[n]{|a_n|} = 1,$$

并不隐含关于级数收敛的任何信息. 这样的一个级数可以收敛也可以发散. 例如，级数

$$\sum_{n=1}^{\infty}\frac{1}{n},$$

对于它 $\lim_{n\to\infty}\sqrt[n]{|a_n|} = 1$ 和 $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = 1$，正像我们在第 539

页中看到的这级数是发散的，另一方面，正像我们即将看到的，级数 $\sum_{n=1}^{\infty}\frac{1}{n^2}$ 满足同样的关系式，但它却是收敛的.

作为我们的判别法的应用的例，我们先考虑级数

$$q+2q^2+3q^3+\cdots+nq^n+\cdots.$$

对于这个级数

$$\lim_{n\to\infty}\sqrt[n]{|a_n|}=|q|\cdot\lim_{n\to\infty}\sqrt[n]{n}=|q|,$$

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|=|q|\cdot\lim_{n\to\infty}\frac{n+1}{n}=|q|.$$

根据比值判别法和根式判别法，甚至较弱的公式(5)，都可以推知，如果 $|q|<1$，则该级数是收敛的.

另一方面，如果我们考虑级数

$$1+2q+q^2+2q^3+\cdots+q^{2n}+2q^{2n+1}+\cdots,$$

当 $\frac{1}{2}\leqslant|q|<1$ 时，我们不再能够根据比值判别法来证明级数的收敛性，因为 $\left|\frac{2q^{2n+1}}{q^{2n}}\right|=2|q|\geqslant 1$. 但是根式判别法立即给出 $\lim_{n\to\infty}\sqrt[n]{|a_n|}=|q|$，因而说明，在 $|q|<1$ 的条件下，级数收敛. 当然，在 $|q|<1$ 的条件下，我们也可以直接观察.

c. 与积分相比较[1)]

我们现在讨论一个研究收敛性的全然不同的方法. 我们将以特别简单而重要的级数

$$\sum_{n=1}^{\infty}\frac{1}{n^\alpha}=1+\frac{1}{2^\alpha}+\frac{1}{3^\alpha}+\cdots$$

作为典型来说明这个方法，式中一般项 a_n 是 $\frac{1}{n^\alpha}$，α 是一个正数. 为了研究这个级数是收敛的还是发散的，我们考察函数 $y=\frac{1}{x^\alpha}$ 的图形，同时在 x 轴上标出整数的横坐标 $x=1$, $x=2,\cdots$. 我们

---

1) 与这相联系也可以参看第六章的附录第 532 页.

首先在 x 轴上的区间 $n-1\leqslant x\leqslant n(n>1)$ 上作高为 $\frac{1}{n^\alpha}$ 的矩形，并且把它与在 x 轴的同一区间上，区间端点的垂直线和曲线 $y=\frac{1}{x^\alpha}$ 所围成的面积(这个范围在图 7.2 中以阴影部分表出)进行比较. 其次，我们在区间 $n\leqslant x\leqslant n+1$ 上作高为 $\frac{1}{n^\alpha}$ 的矩形，类似地把它与在同一个区间上，曲线下围起来的面积(这个范围在图 7.2 中以交叉网格表出)进行比较. 在第一种情形，曲线下的面积

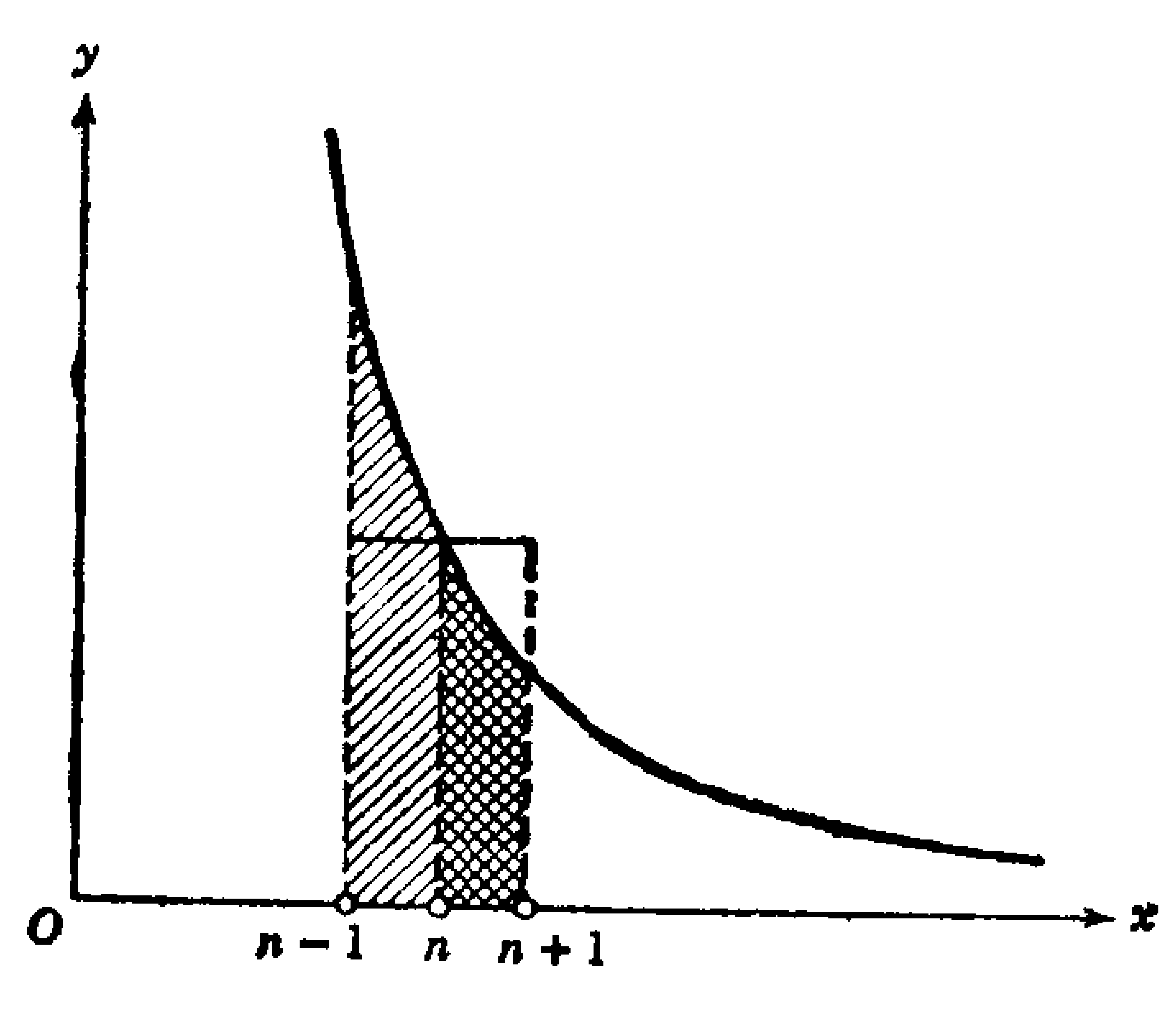


图 7.2 级数与积分的比较

明显地大于矩形的面积；在第二种情形，曲线下的面积小于矩形的面积. 换句话说，

$$\int_{n}^{n+1}\frac{dx}{x^\alpha}<\frac{1}{n^\alpha}<\int_{n-1}^{n}\frac{dx}{x^\alpha}.$$

分别对于 $n=1,2,3,\cdots,m$ 和 $n=2,3,\cdots,m$，写下这些不等式并且求和，我们就得到以下关于第 m 个部分和 $s_m=\sum_{n=1}^{m}\frac{1}{n^\alpha}$ 的估计：

$$\int_{1}^{m+1}\frac{dx}{x^\alpha}<s_m<1+\int_{1}^{m}\frac{dx}{x^\alpha}. \tag{6}$$

现在，当 m 增加时，积分 $\int_{1}^{m}\frac{dx}{x^\alpha}$ 趋向于一个有穷极限还是递增无

极限依 $\alpha > 1$ 还是 $\alpha \leq 1$ 而定. 从而单调数序列 s_n 是有界还是递增无界依 $\alpha > 1$ 还是 $\alpha \leq 1$ 而定，于是我们有下面的定理.

定理 倒数幂的级数

$$\sum_{n=1}^{\infty} \frac{1}{n^{\alpha}} = \frac{1}{1^{\alpha}} + \frac{1}{2^{\alpha}} + \frac{1}{3^{\alpha}} + \cdots$$

当且仅当 $\alpha > 1$ 时收敛.

对于 $\alpha = 1$，我们以前用不同的方法证明过的调和级数的发散性，现在是定理的一个直接的结果；同样，级数

$$\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots$$

$$\frac{1}{1^3} + \frac{1}{2^3} + \frac{1}{3^3} + \cdots$$

收敛而级数 $\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \cdots$ 发散.

对于 $\alpha > 1$ 的收敛级数 $\sum_{\nu=1}^{\infty} \frac{1}{\nu^{\alpha}}$ 在收敛性的研究中经常被用来作为比较级数. 例如，我们立即看出，对于 $\alpha > 1$，级数 $\sum_{\nu=1}^{\infty} \frac{c_\nu}{\nu^{\alpha}}$ 当系数的绝对值 $|c_\nu|$ 小于一个固定的不依赖于 ν 的上界时就绝对收敛.

欧拉常数. 从估计式 (6) 对于 $\alpha = 1$ 立即得出数序列

$$c_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \log n$$

$$= s_n - \log n > \log(n+1) - \log n > 0$$

是下有界的. 因为从不等式

$$\frac{1}{n+1} < \int_n^{n+1} \frac{dx}{x} = \log(n+1) - \log n$$

$$= \frac{1}{n+1} + C_n - C_{n+1},$$

我们看到序列是单调递减的，于是它必定趋向一个极限

$$\lim_{n\to\infty} C_n = \lim_{n\to\infty}\left(1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}-\log n\right)=C.$$

数 C 的值是 $0.5772\cdots$，称为欧拉常数. 与分析中的其它重要的特殊数，例如 π 和 e 不同，对于欧拉常数没有发现另外的具有简单的形成规律的表达式. 到今天为止，还不知道 C 是有理数还是无理数.

7.3 函数序列

正如以前经常强调的，极限过程不仅用于以其他更简单的数近似地表示已知数，而且也用于把已知数的集合扩展成为更广泛的集合.下述事实在分析中具有决定性的意义:不仅要研究常数项序列或常数项无穷级数的极限，而且类似地，也要研究函数项序列或其项是一个变量 x 的函数的级数的极限，如像泰勒级数或一般的幂级数. 不仅用更简单的函数去逼近给定的函数需要这样的极限过程，而且新的函数的定义和分析描述也往往需要建立在函数序列的极限概念的基础上: $f(x)=\lim f_n(x)$ 对 $n\to\infty$. 等价地，我们可以把 $f(x)$ 和 $f_n(x)$ 看作函数 $g_n(x)$ 的无穷级数 $f(x)=\sum_{\nu=1}^{\infty} g_\nu(x)$ 的和与部分和，其中 $n>1$ 时 $g_n(x)=f_n(x)-f_{n-1}(x)$ 和 $g_1(x)=f_1(x)$.

我们现在来讨论确切的定义和几何解释.

a. 函数与曲线序列的极限过程

定义 如果在区间 $a\leqslant x\leqslant b$ 的每一点 x 处，函数值 $f_n(x)$ 在通常意义下收敛到函数值 $f(x)$，则称序列 $f_1(x)$，$f_2(x)$，$\cdots$ 在区间 $a\leqslant x\leqslant b$ 上收敛到极限函数 $f(x)$. 在这种情形下，我们写成 $\lim_{n\to\infty} f_n(x)=f(x)$. 根据柯西判别法(参看第 78 页)，我们能够不涉及极限函数 $f(x)$ 来叙述序列的收敛性: 函数序列收敛到一个极限函数，当且仅当在所考虑的区间里的每一个点 x 处，对于每一个正数 ε，量 $|f_n(x)-f_m(x)|$ 是小于 ε 的，只要 n 和 m 选得充分

大，也就是，大于某一个数 N. 这个数 $N=N(\varepsilon, x)$ 通常是依赖于 ε 和 x 的，并且当 ε 趋于零时无限增大.

我们经常遇见函数序列的极限的情形. 我们仅提一提，当 α 为无理数时幂 x^α 的定义所根据的是方程

$$x^\alpha = \lim_{n\to\infty} x^{r_n},$$

其中 $r_1, r_2, \cdots, r_n, \cdots$ 是一个趋于 α 的有理数的序列；或者再提到方程

$$e^x = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n,$$

其中右边的逼近函数 $f_n(x)$ 是 n 次多项式.

使用曲线来作为函数的图形表示这件事启示我们，也可以谈谈关于曲线序列的极限. 譬如说，前面提到的极限函数 x^α 和 e^x 的图形可以分别看作是函数 x^{r_n} 和 $\left(1+\frac{x}{n}\right)^n$ 的图形的极限曲线.

然而，在函数的极限和曲线的极限之间存在着一个细微的差别，直到十九世纪中叶一直没有被清楚地观察到. 我们将用一个例来说明这一点，然后再在下一节去系统地讨论它.

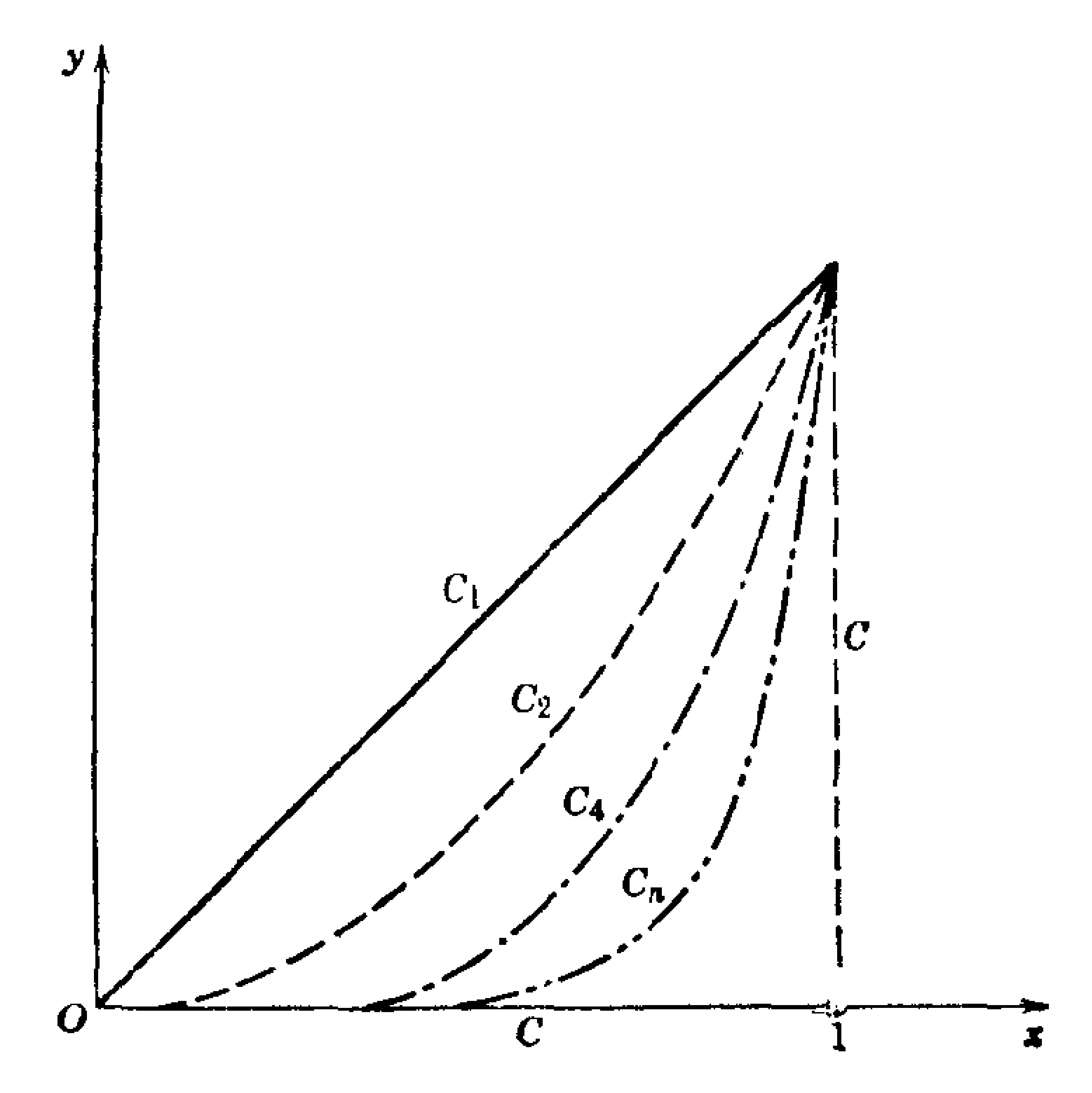


图 7.3　极限曲线与极限函数

我们考虑区间 $0 \leqslant x \leqslant 1$ 上这一序列函数

$$f_n(x) = x^n, \quad n = 1, 2, \cdots.$$

这些函数都是连续的，并且极限函数 $\lim\limits_{n \to \infty} f_n(x) = f(x)$ 存在. 但是这个极限函数是不连续的. 相反，因为对 n 的一切值，函数值 $f_n(1) = 1$，极限

$$f(1) = 1;$$

而另一方面，对于 $0 \leqslant x < 1$，正像我们在第一章第 68 页看到的，极限 $f(x) = \lim\limits_{n \to \infty} f_n(x) = 0$. 所以函数 $f(x)$ 是一个间断函数，在 $x = 1$ 处有值 1，而对于区间上所有其他的 x 值有值 0.

这个间断性在几何上是用函数 $y = f_n(x)$ 的图形 C_n 来说明的. 这些图形(参看第 69 页图 1.44)都是连续曲线，它们都经过原点和 $x = 1$, $y = 1$ 这个点，并且当 n 增大时，它们越来越靠近 x 轴. 这些曲线确实具有一个极限曲线 C，它完全不是间断的，而是由 x 轴上 $x = 0$ 和 $x = 1$ 之间的一段以及直线 $x = 1$ 上 $y = 0$ 和 $y = 1$ 之间的一段所组成的. 于是这些曲线收敛到具有一个垂直部分的连续的极限曲线，而函数则收敛到一个间断的极限函数. 我们由此认识到，极限函数的这个间断性是由极限曲线上有垂直于 x 轴的部分来表明的. 这个极限曲线不是极限函数的图形；因为对应于垂直部分的 x 值，曲线给出无穷多个 y 值，而函数却只给出一个值. 因此，函数 $f_n(x)$ 的图形的极限并不同于这些函数的极限 $f(x)$ 的图形.

当然，相应的叙述对于无穷级数也是一样成立的.

7.4 一致收敛与不一致收敛

a. 一般说明和定义

函数的收敛性概念与曲线的收敛性概念之间的差别是学生应当清晰地把握住的现象. 这包括函数的序列或无穷函数级数的所谓不一致收敛性，对此我们将讨论详细一些.

在区间 $a \leqslant x \leqslant b$ 上，函数 $f(x)$ 是序列函数 $f_1(x), f_2(x), \cdots$ 的极限，按照定义的意思仅仅是，在区间上的每一个点 x 处通常的

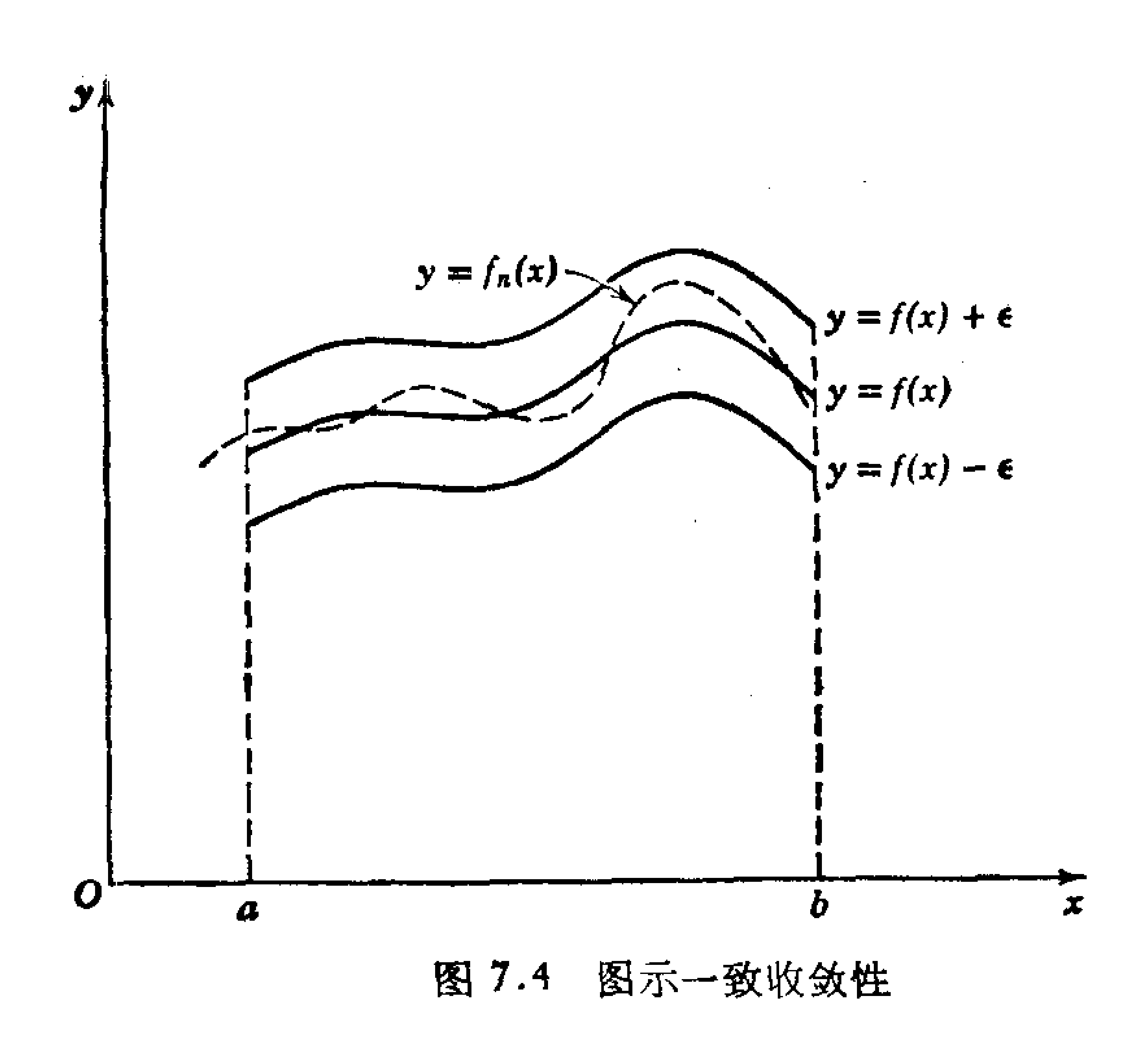


图 7.4 图示一致收敛性

极限关系 $f(x)=\lim\limits_{n\to\infty} f_n(x)$ 成立.

这样的收敛性是序列在点 x 处的一种局部性质. 然而,自然可以要求比上述的逼近仅是局部收敛更高一些,即,如果我们指定一个任意的精确度 ε, 则从某一个指标数 N 以后, 对于所有的 x 值, 所有的函数 $f_n(x)$ 应该界于 $f(x)-\varepsilon$ 和 $f(x)+\varepsilon$ 之间, 以致 $y=f_n(x)$ 的图形全都落在图 7.4 所标明的窄条里. 如果能够使得逼近的精确度同时在区间上处处至少等于预先指定的一个正数 ε, 也就是,处处选出了同一个不依赖于 x 的数, $N(\varepsilon)$ 我们就说逼近是一致的[1]. 如果 $\lim\limits_{n\to\infty} f_n(x)=f(x)$ 对于 $a\leqslant x\leqslant b$ 一致地成立, 则对于每一个 $\varepsilon>0$, 存在相应的数 $N=N(\varepsilon)$, 使得 $|f(x)-f_n(x)|<\varepsilon$ 对所有的 $n>N$ 和区间上所有的 x 都成立. 在十九世纪中叶, 当塞德尔和其他的人提醒人们注意函数的收敛性完全不一定是一致的, 决不能天真地认为函数收敛就一致收敛时,使许多人大为震惊.

不一致收敛的例. 一致收敛的概念是从不一致收敛的各种例来得到说明

1) 与第 42 页上一致连续性的类似的定义相比较, 在那里我们能够选取同一个不依赖于 x 的数 $\delta(\varepsilon)$.

的.

(a) 第一个例是刚才考察过的函数序列，$f_n(x)=x^n$；在区间 $0\leqslant x\leqslant 1$ 上，这个序列收敛到极限函数 $f(x)$，当 $0\leqslant x<1$ 时 $f(x)=0$，而 $f(1)=1$. 在区间的每一点处都是收敛的；就是说，如果 ε 是任何一个正数，并且如果我们选择任何一个确定的值 $x=\xi$，则当 n 充分大时不等式 $|\xi^n-f(\xi)|<\varepsilon$ 是确定地成立的. 然而，这个逼近却不是一致的. 因为，如果我们选择 $\varepsilon=\frac{1}{2}$，于是数 n 不论选得多么大，我们总能够找到一个点 $x=\eta\neq 1$，在这个点处 $|\eta^n-f(\eta)|=\eta^n>\frac{1}{2}$；事实上，这是对所有的点 $x=\eta$，其中 $1>\eta>\sqrt[n]{\frac{1}{2}}$ 都成立的. 因此不可能选到一个如此大的数 n 使得对整个区间上全体的点 $f(x)$ 和 $f_n(x)$ 之间的差都小于$\frac{1}{2}$.

如果我们参考这些函数的图形 (图 7.3)，这一状况就变得清楚明白了. 我们看到，无论我们选择多么大的一个 n 值，对于仅比 1 小一点的 ξ 值，函数 $f_n(\xi)$ 的值将非常靠近 1，因此不能够很好地逼近于值为 0 的 $f(\xi)$.

类似的情况由函数

$$f_n(x)=\frac{1}{1+x^{2n}}$$

在点 $x=1$ 和点 $x=-1$ 的邻域内展示出来；这是容易确立的. 在这里当 $|x|<1$ 时 $f(x)=1$，当 $|x|=1$ 时 $f(x)=\frac{1}{2}$，而当 $|x|>1$ 时 $f(x)=0$.

(b) 在上面两个例里，收敛的不一致性是与极限函数不连续这个事实有关的. 然而也容易构造一个连续函数的序列，它们确实是收敛到一个连续的极限函数，但不是一致收敛的. 我们仅注意区间 $0\leqslant x\leqslant 1$，同时对于 $n\geqslant 2$ 给出下面的定义：

$$f_n(x)=xn^\alpha \qquad 当 \qquad 0\leqslant x\leqslant\frac{1}{n},$$

$$f_n(x)=\left(\frac{2}{n}-x\right)n^\alpha \qquad 当 \qquad \frac{1}{n}\leqslant x\leqslant\frac{2}{n},$$

$$f_n(x)=0 \qquad 当 \qquad \frac{2}{n}\leqslant x\leqslant 1,$$

这里在开始的时候，我们能够对 α 选取任何值，但是一经选定之后，对于序列的所有项这个 α 值就必须保持固定. 从图形上看，我们的函数由 x 轴上的区间 $0\leqslant x\leqslant\frac{2}{n}$ 上的两根线段所组成的屋顶形状所表示，而从 $x=\frac{2}{n}$ 以后

图形是 x 轴自己(参看图 7.5).

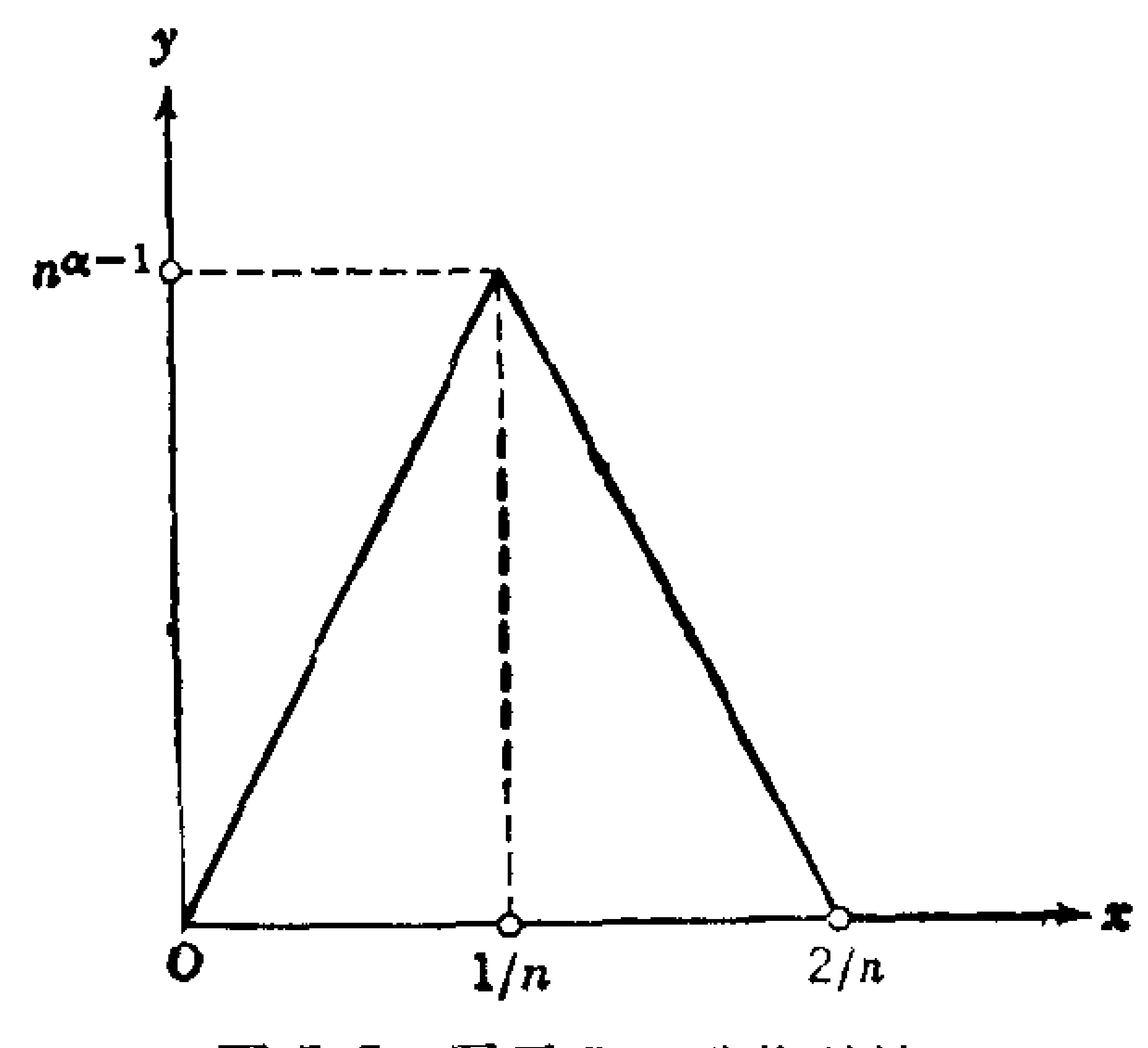


图 7.5　图示非一致收敛性

如果 $\alpha<1$，图形最高点的高度一般地有值 $n^{\alpha-1}$，当 n 增大时，将趋于零；于是曲线就将趋向 x 轴，因而函数 $f_n(x)$ 将一致地收敛到极限函数 $f(x)=0$.

如果 $\alpha=1$，对于每一个 n 值,图形的顶端将有高度1. 如果 $\alpha<1$，当 n 增大时,图形顶端的高度将无限增大.

然而，无论 α 如何选取，序列 $f_1(x)$，$f_2(x)$，$\cdots$ 总是趋向于极限函数 $f(x)=0$. 因为,假设 x 是正的,则对于所有充分大的 n 值我们有 $2/n<x$,从而 x 不在图形的尖顶部分下面，并且 $f_n(x)=0$；当 $x=0$ 时,所有的函数值 $f_n(x)$ 等于 0，因而不论哪种情况都有 $\lim\limits_{n\to\infty}f_n(x)=0$.

然而，如果 $\alpha\geqslant1$，收敛一定是不一致的；因为显然不可能选到足够大的 n 使得表达式 $|f(x)-f_n(x)|=f_n(x)$ 在区间上处处小于 $\frac{1}{2}$.

(c) 完全地类似的情况由函数序列

$$f_n(x)=xn^\alpha e^{-nx}$$

显示出来，与前面的情形对比，其中序列的每一个函数是被一个单一的分析式子表示出来. 这里方程 $\lim\limits_{n\to\infty}f_n(x)=0$ 再一次对每一个正的 x 值成立，因为当 n 增大时,函数 e^{-nx} 趋于零的数量级大于 $\frac{1}{n}$ 的任何次幂(参看第 266 页 3.7b). 当 $x=0$ 时,我们总有 $f_n(x)=0$，因此对于区间 $0\leqslant x\leqslant a$ 上的每一个 x 值,其中 a 是任意一个正数,都有

$$f(x)=\lim_{n\to\infty}f_n(x)=0,$$

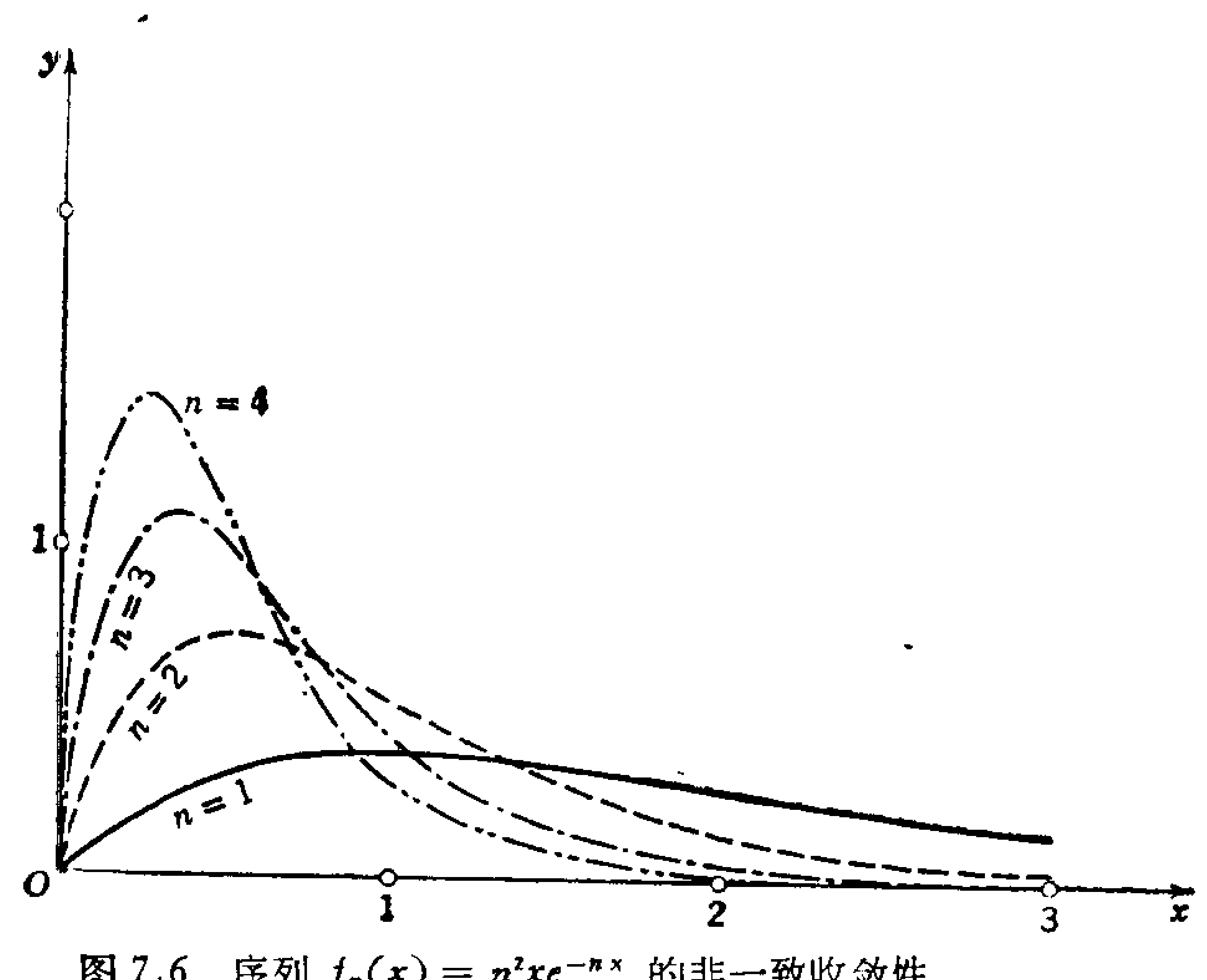


图 7.6　序列 $f_n(x)=n^2xe^{-nx}$ 的非一致收敛性

但这里又一次显出不是一致地收敛到极限函数. 因为在 $x=\frac{1}{n}$ 的点上(在这个点上 $f_n(x)$ 有它的最大值),我们有

$$f_n(x)=f_n\left(\frac{1}{n}\right)=\frac{n^{\alpha-1}}{e},$$

于是我们看出,如果 $\alpha\geqslant 1$, 收敛是不一致的,因为对于每一条曲线 $y=f_n(x)$, 无论 n 选得多么大,总会包含这样的点(即, $x=\frac{1}{n}$ 的点, 这个点是随 n 变化的),在这个点上 $f_n(x)-f(x)=f_n(x)>\frac{1}{2e}$ (参看图 7.6).

(d) 一致收敛和不一致收敛的概念，当然可以被推广到无穷级数的情形. 我们说一个级数

$$g_1(x)+g_2(x)+\cdots$$

是一致收敛的或不一致收敛的,是按照它的部分和 $f_n(x)$ 的情形而言. 一个很简单的不一致收敛级数的例是

$$f(x)=x^2+\frac{x^2}{1+x^2}+\frac{x^2}{(1+x^2)^2}+\frac{x^2}{(1+x^2)^3}+\cdots.$$

当 $x=0$ 时, 每一个部分和 $f_n(x)=x^2+\cdots+x^2/(1+x^2)^{n-1}$ 都有值为 0; 因此 $f(0)=0$. 当 $x\neq 0$ 时,级数简单地是一个几何级数,具有正的比 $1/(1+x^2)<1$; 因此我们能够按照初等法则来求和,这样对于每一个 $x\neq 0$

就得到和

$$\frac{x^2}{1-1/(1+x^2)}=1+x^2.$$

于是极限函数 $f(x)$ 除了在 $x=0$ 点，处处由表达式 $f(x)=1+x^2$ 给出，而 $f(0)=0$；因此在原点有一个可去的不连续点.

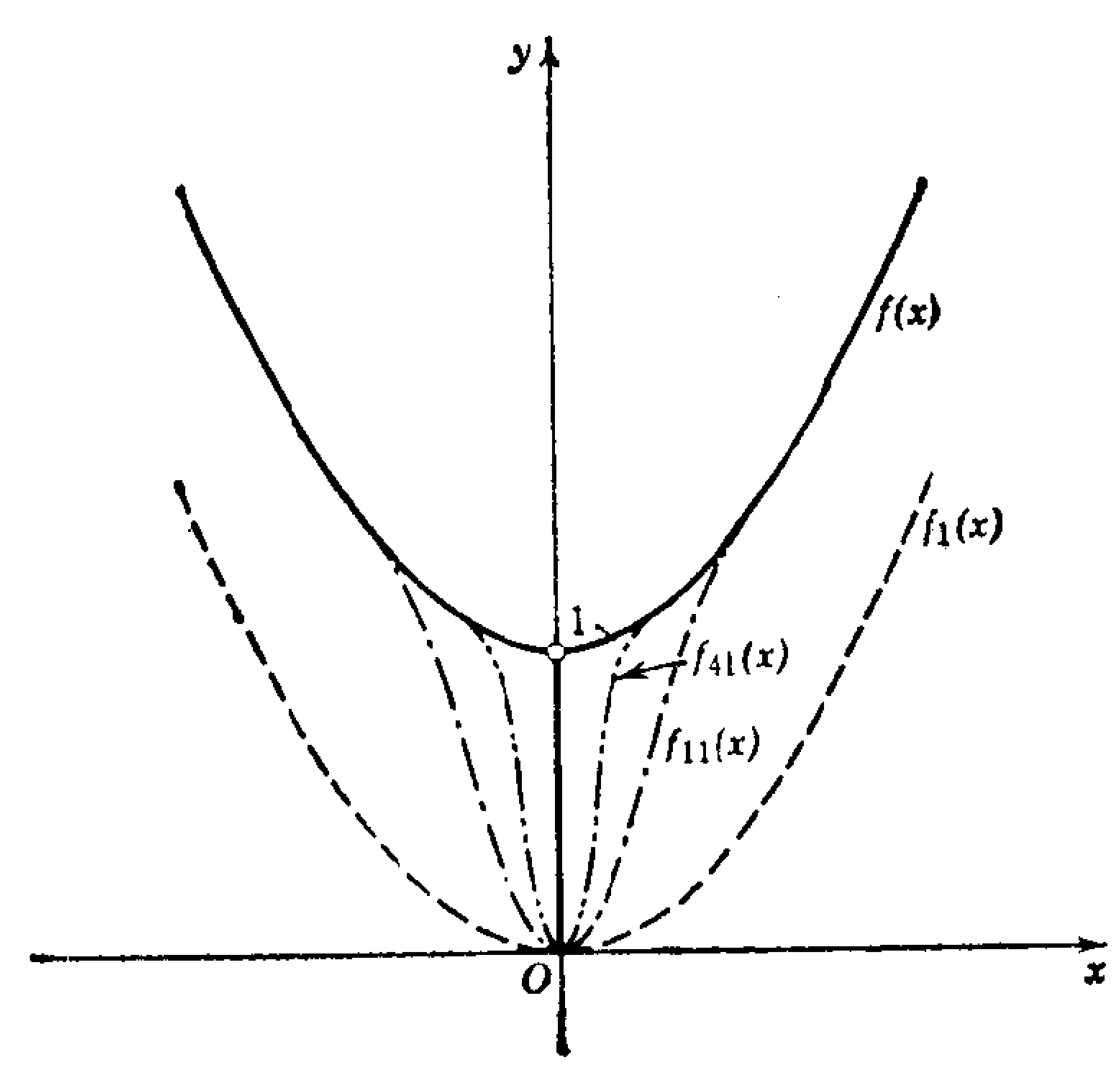


图 7.7　收敛到具有可去跳跃间断点的函数

这里在每一个包含原点的区间上，我们又一次得到不一致收敛性. 因为当 $x=0$ 时，差 $f(x)-f_n(x)=r_n(x)$ 总是为 0，而在 x 的所有其他的点上 r_n 由表达式 $r_n(x)=1/(1+x^2)^{n-1}$ 给出，读者可以自己验证这一点. 如果我们需要这个表达式小于，譬如说是 $\frac{1}{2}$，那末对于每个固定的 x 值，这是能够由选择足够大的 n 来实现的. 但是我们不能够找到充分大的 n 值去保证 $r_n(x)$ 处处小于 $\frac{1}{2}$；因为即使我们选择了不论多么大的 n，我们仍能够在足够靠近 0 的地方选取 x 使 $r_n(x)$ 大于 $\frac{1}{2}$. 因此不超过 $\frac{1}{2}$ 的一致逼近是不可能的. 如果我们考察一下逼近曲线（参看图 7.7），事情就变得很清楚了. 这些曲线除了在 $x=0$ 点的附近外，当 n 增大时，是越来越靠近抛物线 $y=1+x^2$ 的；然而，在 $x=0$ 点附近，这些曲线越来越窄地往下延伸到原点，并且当 n 增大时，这个延伸越来越紧靠一条确定的直线段，即 y 轴的一部分，因此我们得到的极限曲线是一条抛物线加上一段垂直向下延伸到原点的直线

段.

作为非一致收敛的一个进一步的例,我们提出这一级数$\sum_{\nu=0}^{\infty} g_\nu(x)$，其中$g_\nu(x)$定义在区间$0 \leqslant x \leqslant 1$上，当$\nu \geqslant 1$时$g_\nu(x)=x^\nu-x^{\nu-1}$，$g_0(x)=1$. 这个级数的部分和是函数$x^\nu$，这已经在第557页例(a)中考察过.

b. 一致收敛的一个判别法

前面的考察告诉我们，一个序列或级数的一致收敛性是一个特殊的性质，它并不为所有的序列和级数所具备. 我们现在对于无穷级数,重复叙述一致收敛的定义：如果$f(x)$能够由一个固定的但项数充分大的和$g_1(x)+\cdots+g_N(x)=f_N(x)$在区间上与$x$无关地逼近到近似界限$\varepsilon$之内(其中$\varepsilon$是一个任意小的正数)，则称级数

$$g_1(x)+g_2(x)+\cdots$$

在这个区间上一致收敛到函数$f(x)$.

我们又有一个不需要知道极限函数$f(x)$的一致收敛的判别法(柯西判别法)：级数一致地收敛(或等价地说，函数序列$f_n(x)$一致地收敛)当且仅当n和m大于一个不依赖于x的数N时,差$|f_n(x)-f_m(x)|$能够在区间上处处小于一个任意的量ε. 首先因为,如果收敛是一致的,我们能够选取n和m大于一个不依赖于x的数N,使得$|f_n(x)-f(x)|$和$|f_m(x)-f(x)|$同时小于$\varepsilon/2$，从这就得到$|f_n(x)-f_m(x)|<\varepsilon$；其次,如果当$n$和$m$大于$N$时,对所有的$x$值$|f_n(x)-f_m(x)|<\varepsilon$，那末在选取的$n>N$的任意固定值上,让$m$趋于无穷大,对于每一个$x$值,我们有关系式

$$|f_n(x)-f(x)|=\lim_{m\to\infty}|f_n(x)-f_m(x)|\leqslant\varepsilon,$$

因此收敛是一致的.

正像我们将要看到的，正是这个一致收敛的条件使无穷级数和其他的作用于函数的极限过程成为分析的方便有用的工具. 幸运的是,在分析和它的应用里经常遇到的极限过程中,非一致收敛仅仅在孤立的例外点上出现,因而在目前不致于困扰我们.

一个级数的一致收敛性通常是由下面的准则确立的(把它与

常数项控制级数进行比较):

如果级数 $\sum_{\nu=1}^{\infty} g_\nu(x)$ 的项满足条件 $|g_\nu(x)| \leqslant a_\nu$，其中数 a_ν 是正的常数并且形成一个收敛级数 $\sum_{\nu=1}^{\infty} a_\nu$，那末级数 $\sum_{\nu=1}^{\infty} g_\nu(x)$ 一致(并且绝对地)收敛.

因为这时我们有

$$\left|\sum_{\nu=n}^{m} g_\nu(x)\right| \leqslant \sum_{\nu=n}^{m}\left|g_\nu(x)\right| \leqslant \sum_{\nu=n}^{m} a_\nu,$$

而且因为根据柯西判别法，当 n 和 $m > n$ 选得足够大时，和 $\sum_{\nu=n}^{m} a_\nu$ 能够成为任意小，这确切地表达了关于一致收敛性的必要充分条件.

第一个例是由几何级数 $1 + x + x^2 + \cdots$ 提供的，其中 x 被限制在区间 $|x| \leqslant q$ 上，q 是任意的小于 1 的正数. 于是级数的项在数量上小于或等于收敛的几何级数 $\sum q^\nu$ 的项.

进一步的例是由"三角级数"

$$\frac{c_1 \sin(x-\delta_1)}{1^2} + \frac{c_2 \sin(x-\delta_2)}{2^2} + \frac{c_3 \sin(x-\delta_3)}{3^2} + \cdots$$

给出，以 $|c_n| < c$ 为条件，其中 c 是一个不依赖于 n 的正常数. 因为这时我们有

$$g_n(x) = \frac{c_n \sin(x-\delta_n)}{n^2}, \quad 于是 \quad |g_n(x)| < \frac{c}{n^2}.$$

因此根据级数 $\sum_{n=1}^{\infty} \frac{c}{n^2}$ 的收敛性立即得出三角级数的一致与绝对收敛性.

c. 连续函数的一致收敛级数之和的连续性

一致收敛的重要性在于，一个一致收敛的级数在许多方面的情况确实很像有穷个函数的和. 举例说罢，有穷个连续函数的和仍是连续的，相应地，我们有下面的定理.

定理 如果一个连续项的级数在一个区间里一致地收敛，则它的和也是一个连续函数.

证明. 证明是十分简单的. 我们把级数

$$f(x) = g_1(x) + g_2(x) + \cdots$$

分成第 n 个部分和加上余项 $R_n(x)$. 像通常一样，$f_n(x) = g_1(x) + \cdots + g_n(x)$. 现在如果指定了任意的正数 ε，我们就能够根据一致收敛性选取充分大的数 n 使得在整个区间上余项都小于 $\varepsilon/4$. 因此

$$|R_n(x+h) - R_n(x)| < \frac{\varepsilon}{2}$$

对区间上每一对数 x 和 $x+h$ 都成立. 部分和 $f_n(x)$ 是由有穷个连续函数的和组成的，因而是连续的了，于是对于区间上的每一个点 x，我们能够选取充分小的一个正数 δ，使得当 $|h| < \delta$ 且点 x 和 $x+h$ 都属于区间时，

$$|f_n(x+h) - f_n(x)| < \frac{\varepsilon}{2}$$

成立. 于是得到

$$\begin{aligned}|f(x+h)-f(x)| &= |f_n(x+h)-f_n(x)+R_n(x+h)-R_n(x)| \\ &\leqslant |f_n(x+h)-f_n(x)| + |R_n(x+h)-R_n(x)| < \varepsilon,\end{aligned}$$

这表示了我们的函数的连续性.

当我们从前面的各个例中，回忆起连续函数的不一致收敛级数的和不一定是连续的这个论述时，这个定理的重要性就成为明显的了.从上述定理我们可以得出结论:如果连续函数的收敛级数的和有一个不连续点，那末在这个点的每一个邻域里收敛都是不一致的. 因此，用连续函数的级数去表示不连续函数的任何一种表示法，都必然建立在利用不一致收敛极限过程的基础上.

d. 一致收敛级数的积分

有穷个连续函数的和能够“逐项”积分;即，和的积分由每一项分别积分以后相加得到. 在连续函数的收敛无穷级数的情况下，只要在积分区间上级数一致收敛，同样可以逐项积分.

在一个区间上一致收敛的级数 $\sum_{\nu=1}^{\infty} g_\nu(x) = f(x)$ 在这个区间

上能够逐项进行积分；或者，说得更明白些，如果 a 和 x 是一致收敛区间上的两个数，那末级数 $\sum_{\nu=1}^{\infty}\int_a^x g_\nu(t)\,dt$ 收敛，并且事实上关于 x 一致收敛，它的和等于 $\int_a^x f(t)\,dt$[1].

为了证明这一定理，我们和以前一样将 $f(x)$ 写成

$$f(x)=\sum_{\nu=1}^{\infty} g_\nu(x)=f_n(x)+R_n(x).$$

我们已经假定了级数的各项是连续的；因此根据 7.4c 节，和也是连续的，因而是可积的．现在如果 ε 是任意的正数，我们能够找到一个充分大的数 N，使得对区间上每一个 x 值，当每一个 $n>N$ 时，不等式 $|R_n|<\varepsilon$ 成立．根据积分学的中值定理，我们有

$$\left|\int_a^x [f(t)-f_n(t)]\,dt\right|\leqslant \varepsilon l,$$

其中 l 是积分区间的长度．由于有穷和 $f_n(x)$ 的积分能够逐项进行，这就给出

$$\left|\int_a^x f(t)\,dt-\sum_1^n\int_a^x g_\nu(t)\,dt\right|<\varepsilon l.$$

但是由于 εl 要多小就能多小，这就指出了

$$\sum_{\nu=1}^{\infty}\int_a^x g_\nu(t)\,dt=\lim_{n\to\infty}\sum_{\nu=1}^{n}\int_a^x g_\nu(t)\,dt=\int_a^x f(t)\,dt,$$

定理得证．

如果，代替无穷级数，我们希望谈论函数序列，我们的结果就可以用下面的形式来表示：

如果在一个区间上函数序列 $f_1(x)$，$f_2(x)$，$\cdots$ 一致地趋向极

---

1) 注意，在这个定理里我们必须取定积分．举例来说，级数 $\sum_{\nu=1}^{\infty} g_\nu(x)$ 当 $g_\nu(x)=0$ 时一致收敛；然而每一项取不定积分 $\int g_\nu(x)dx=$ 常数 $=c$，则引出一般是发散的级数 $\sum_{\nu=1}^{\infty} c$.

限函数 $f(x)$，那末等式

$$\int_a^b f(x)\,dx = \lim_{n\to\infty}\int_a^b f_n(x)\,dx \tag{7}$$

对属于区间的每一对数 a，b 都成立；换句话说，这时我们能够交换积分运算与极限运算的次序.

这个事实不是微不足道的. 从一个如像在十九世纪曾经流行过的天真的观点看来，这两个过程的可交换性几乎是不容怀疑的；但是看一看 7.4a 里的例，就会告诉我们，在不一致收敛的情况下，上面的等式可以不成立. 我们只需考察第 557 页的例 b 就可以了，在该例中极限函数的积分是 0，而函数 $f_n(x)$ 在区间 $0\leqslant x\leqslant 1$ 上的积分，也就是说，在图 7.5 里三角形的面积，有值

$$\int_0^1 f_n(x)dx = n^{\alpha-2},$$

而当 $\alpha\geqslant 2$ 时这个积分值是不趋向于零的. 在这里，我们直接从图形看到，$\int_0^1 f(x)dx$ 与 $\lim\limits_{n\to\infty}\int_0^1 f_n(x)dx$ 之间的差异就在于收敛的不一致性.

另一方面，当考虑 α 的这样的值 $1\leqslant\alpha\leqslant 2$，我们看到，虽然收敛是不一致的，等式 $\lim\limits_{n\to\infty}\int_0^1 f_n(x)dx=\int_0^1 f(x)dx$ 却能够成立. 作为进一步的例，我们有级数 $\sum\limits_0^\infty g_n(x)$，其中当 $n\geqslant 1$ 时 $g_n(x)=x^n-x^{n-1}$ 而 $g_0(x)=1$，能够在 0 到 1 之间逐项积分，虽然它不是一致收敛的. 因此对于逐项可积来说，虽然收敛的一致性是一个充分条件，但它决不是一个必要条件.

e. 无穷级数的微分法

一致收敛级数或序列关于微分的情形与积分是完全不同的. 例如，函数序列 $f_n(x)=\dfrac{\sin n^2x}{n}$ 确实一致收敛到极限函数 $f(x)=0$，但是导数 $f_n'(x)=n\cos n^2x$ 确实不处处收敛到极限函数的导数 $f'(x)=0$，当我们考虑 $x=0$ 时就可以看到. 因此，即使在一致收敛的情况下，我们也不能交换微分过程与极限过程的次序.

当然相应的叙述对于无穷级数也是成立的. 例如，级数

$$\sin x+\frac{\sin 2^4x}{2^2}+\frac{\sin 3^4x}{3^2}+\cdots$$

是绝对和一致收敛的，因为在数值上它的项不大于收敛级数 $\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\cdots$ 相应的项. 然而，如果逐项微分这个级数，我们就得到级数

$$\cos x + 2^2\cos 2^4 x + 3^2\cos 3^4 x + \cdots,$$

显然，这个级数在 $x=0$ 发散.

保证在特殊情况下可以逐项微分的唯一有用的准则，由下面的定理给出。

如果逐项微分一个收敛的无穷级数 $\sum_{\nu=0}^{\infty} G_\nu(x)=F(x)$ 以后，我们得到一个具有连续的项的一致收敛级数 $\sum_{\nu=0}^{\infty} g_\nu(x)=f(x)$，那么最后这个级数的和就等于第一个级数的和的导数.

因此这个定理特别要求，在逐项微分这个级数以后，我们还必须研究微分的结果是否是一个一致收敛的级数.

这个定理的证明几乎是浅显的. 因为根据 7.4d 节的定理，我们能够逐项积分由微分法得到的这个级数.回忆 $g_\nu(t)=G'_\nu(t)$，我们得到

$$\int_a^x f(t)dt=\int_a^x\left(\sum_{\nu=0}^{\infty} g_\nu(t)\right)dt=\sum_{\nu=0}^{\infty}\int_a^x g_\nu(t)dt$$

$$=\sum_{\nu=0}^{\infty}(G_\nu(x)-G_\nu(a))=F(x)-F(a).$$

对于一致收敛区间里的每一个 x 值这都是成立的，由此得出

$$f(x)=F'(x),$$

于是定理得证.

7.5 幂 级 数

在无穷级数中**幂级数**占有最重要的地位. 所谓幂级数我们指的是这样类型的级数

$$P(x)=c_0+c_1x+c_2x^2+\cdots=\sum_{\nu=0}^{\infty} c_\nu x^\nu \tag{8}$$

（"x 的幂级数"），或更一般地

$$P(x) = c_0 + c_1(x - x_0) + c_2(x - x_0)^2 + \cdots$$

$$= \sum_{\nu=0}^{\infty} c_\nu (x - x_0)^\nu \tag{8a}$$

（"$(x - x_0)$ 的幂级数"），其中 x_0 是一个固定的数．如果在后面这个级数里我们引进 $\xi = x - x_0$ 作为一个新的变量，它就成为新变量 ξ 的一个幂级数 $\sum_{\nu=0}^{\infty} c_\nu \xi^\nu$，因此我们可以仅限于讨论较特殊形式的幂级数 $\sum_{\nu=0}^{\infty} c_\nu x^\nu$ 而不失去一般性．

在第五章（第 470 页）我们考察了函数由多项式近似表示的问题，并从而引出了函数的泰勒级数展开式；而事实上，这泰勒级数就是幂级数．在这一节我们将更为详细地研究幂级数，并将比以前更方便地得到一些最重要的函数的级数展开式．

a．幂级数的收敛性质——收敛区间

存在这样的幂级数，它不对 x 的任何值收敛，当然除了 $x=0$，例如，级数

$$x + 2^2x^2 + 3^3x^3 + \cdots + n^nx^n + \cdots.$$

因为如果 $x \neq 0$，我们能够找到一个整数 N 使得 $|x| > \frac{1}{N}$．于是当 $n > N$ 时，所有的项 n^nx^n 的绝对值将大于 1，而且事实上，当 n 增加时，n^nx^n 将无限增大，因此级数不收敛．

另一方面，存在这样的级数，它对 x 的每一个值都收敛，例如，指数函数的幂级数

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots,$$

它对 x 的每一个值的收敛性可从比值判别法（第 548 页准则 5a）立即得到．第 $n+1$ 项被第 n 项除给出 x/n，因而，无论数 x 怎样选，当 n 增加时，这个比值总趋于零．

下面的基本定理表明了幂级数关于收敛性的情况．

如果一个 x 的幂级数对值 $x=\xi$ 收敛，则它对每一个使得 $|x|<|\xi|$ 的 x 值都绝对收敛，同时在每一个区间 $|x|\leqslant\eta$ 上，级数是一致收敛的，其中 η 是任何一个小于 $|\xi|$ 的正数，这里 η 我们愿意多么靠近 $|\xi|$ 就可以多么靠近.

证明是简单的. 如果级数 $\sum_{\nu=0}^{\infty} c_\nu\xi^\nu$ 收敛，当 ν 增加时，它的一般项趋于零. 由此得到较弱的断言：所有的项都不超过一个与 ν 无关的上界 M，即 $|c_\nu\xi^\nu|<M$. 现在，如果 q 是任意一个满足 $0<q<1$ 的数，并且如果我们把 x 限制在区间 $|x|\leqslant q|\xi|$ 上，就有 $|c_\nu x^\nu|\leqslant|c_\nu\xi^\nu|q^\nu<Mq^\nu$. 因此在这个区间里，我们的级数 $\sum_0^\infty c_\nu x^\nu$ 中的各项的绝对值都小于收敛几何级数 ΣMq^ν 的对应项. 因此从第 562 页的定理得出，级数在区间 $-q|\xi|\leqslant x\leqslant q|\xi|$ 上绝对并一致收敛.

如果一个幂级数不是处处收敛，即，如果存在一个值 $x=\xi$，对于这个值级数发散，则对于每一个使得 $|x|>|\xi|$ 的 x 值级数必定发散. 因为如果对于这样的 x 值级数收敛，根据上面的定理对于数值上比 x 小的 ξ 值，级数就必须收敛.

从这里我们认识到，如果一个幂级数至少对一个不等于零的 x 值收敛，并且至少对一个 x 值发散，那末它就有一个**收敛区间**；即，存在一个确定的正数 ρ 使得当 $|x|>\rho$ 时级数发散，而当 $|x|<\rho$ 时级数收敛. 当 $|x|=\rho$ 时不能得出一般结论. 这里 ρ 正是使级数收敛的 x 值的**最小上界**（这样一个上界的存在是根据第 103 页的定理，因为使得级数收敛的 x 值形成一个有界集合）. 级数只在 $x=0$ 处收敛和级数处处都收敛这两种极端的情形，分别被象征性地表示为 $\rho=0$ 和 $\rho=\infty$.[1]

---

1) 从级数的系数 c_ν 可以直接找到收敛区间. 如果极限 $\lim_{n\to\infty}\sqrt[n]{|c_n|}$ 存在，则

$$\rho=\frac{1}{\overline{\lim}_{n\to\infty}\sqrt[n]{|c_n|}}.$$

对于一般情形，参看第 597 页问题 8.

例如，对于几何级数 $1+x+x^2+\cdots$ 我们有 $\rho=1$；在收敛区间的端点上级数发散．类似地，对于反正切的级数(第 468 页)，

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - + \cdots,$$

我们有 $\rho=1$，从莱布尼兹判别法(第 541 页)我们立即认识到，在收敛区间的两个端点 $x=\pm1$ 处级数都收敛．

由一致收敛性，我们推导出下列的重要事实：在收敛区间内(如果这样一个区间存在的话)，幂级数代表一个连续函数．

b. 幂级数的积分法和微分法

由于一致收敛性，一个幂级数

$$f(x)=\sum_{\nu=0}^{\infty} c_\nu x^\nu$$

在完全位于收敛区间内的任意闭区间上总是允许逐项积分的．于是我们就得到函数

$$F(x)=c+\sum_{\nu=0}^{\infty}\frac{c_\nu}{\nu+1}x^{\nu+1}, \tag{9}$$

对于这个函数 $F'(x)=f(x)$ 并且 $F(0)=c$．

我们也可以在一个幂级数的收敛区间内逐项微分这幂级数，从而得到等式

$$f'(x)=\sum_{\nu=1}^{\infty}\nu c_\nu x^{\nu-1}. \tag{10}$$

为了证明这个断言，我们只需说明，如果把 x 限制在一个完全位于收敛区间内的闭区间上时，右边的级数是一致收敛的．于是假设 ξ 是一个数，我们愿意它多么靠近 ρ 就多么靠近，对于这个 ξ，$\sum_{\nu=1}^{\infty} c_\nu\xi^\nu$ 收敛；于是，正像我们以前所见，数 $|c_\nu\xi^\nu|$ 都不能超过一个与 ν 无关的上界 M，因此 $|c_\nu\xi^{\nu-1}|<\dfrac{M}{|\xi|}=N$ 现在设 q 是任意一个满足 $0<q<1$ 的数；如果我们把 x 限制在区间 $|x|\leqslant q|\xi|$ 上，无穷级数(10)中的各项都不大于级数 $\sum_{\nu=1}^{\infty}|\nu c_\nu q^{\nu-1}\xi^{\nu-1}|$ 中各

对应的项，因此小于级数 $\sum_{\nu=1}^{\infty} N\nu q^{\nu-1}$ 中各对应的项．然而在最后这个级数里，第 $n+1$ 项与第 n 项的比值是 $q(n+1)/n$，它当 n 递增时趋于 q．由于 $0<q<1$，由此[准则(5a)]推知这个级数收敛．因此经过求导得到的级数一致收敛，并且根据第 566 页的定理它表示函数 $f(x)$ 的导数 $f'(x)$，这就证明了我们的断言．

如果我们对幂级数

$$f'(x)=\sum_{\nu=1}^{\infty} \nu c_\nu x^{\nu-1},$$

再一次应用这个结果，经逐项微分就得到

$$f''(x)=\sum_{\nu=2}^{\infty} \nu(\nu-1)c_\nu x^{\nu-2},$$

继续进行这个过程，我们就得出定理：每一个由幂级数表示的函数在收敛区间内可微分任意次，并且可以逐项微分.[1)]

c. 幂级数的运算

前面关于幂级数的情况的定理是我们用与多项式相同的方式对幂级数进行运算的依据.显然，两个幂级数相加或相减，能够由相加或相减对应的系数得到（参看第 546 页）．这也是显然的，一个收敛的幂级数乘上一个常数因子，像任何其他的收敛级数一样，能够由其中每一项乘上该常数因子得到．另一方面，两个幂级数的乘和除却需要较详细地研究，对于这些请读者参考附录（第 582 页）．这里我们(不给证明)仅提到两个幂级数

---

1) 作为 k 阶导数的一个明确的表达式，我们得到

$$f^{(k)}(x)=\sum_{\nu=k}^{\infty} \nu(\nu-1)\cdots(\nu-k+1)c_\nu x^{\nu-k},$$

或者稍微不同的形式

$$\frac{f^{(k)}(x)}{k!}=\sum_{\nu=k}^{\infty}\binom{\nu}{k}c_\nu x^{\nu-k}=\sum_{\nu=0}^{\infty}\binom{k+\nu}{k}c_{k+\nu}x^{\nu}.$$

这两个公式是常用的.

$$f(x)=\sum_{\nu=0}^{\infty}a_\nu x^\nu$$

和

$$g(x)=\sum_{\nu=0}^{\infty}b_\nu x^\nu$$

能够像多项式一样乘在一起. 确切地说，我们有下面的定理：在两个级数的收敛区间的公共部分内，它们的乘积是由收敛幂级数 $\sum_{\nu=0}^{\infty}c_\nu x^\nu$ 给出，其中系数 c_ν 由下列公式给出：

$$\begin{aligned}
c_0 &= a_0b_0,\\
c_1 &= a_0b_1+a_1b_0,\\
c_2 &= a_0b_2+a_1b_1+a_2b_0,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
c_n &= a_0b_n+a_1b_{n-1}+\cdots+a_nb_0,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned}$$

d. 展开式的唯一性

在幂级数的理论里下面的事实是重要的： 如果两个幂级数 $\sum_{\nu=0}^{\infty}a_\nu x^\nu$ 和 $\sum_{\nu=0}^{\infty}b_\nu x^\nu$ 都在包含 $x=0$ 在其内的一个区间里收敛，并且，如果在那个区间里这两个级数表示同一个函数 $f(x)$，则它们是恒等的，即等式 $a_n=b_n$ 对每一个 n 值成立. 换句话说：

一个函数 $f(x)$ 只要能表示成 x 的幂级数，它就只能用一种方式表示成 x 的幂级数.

简短地说：一个函数由幂级数表示的表示法是"唯一的".

为了证明这一定理，我们只需注意到两个幂级数的差，即，具有系数 $c_\nu=a_\nu-b_\nu$ 的幂级数 $\phi(x)=\sum_{\nu=0}^{\infty}c_\nu x^\nu$，在这个区间上表示函数

$$\phi(x)=f(x)-f(x)=0;$$

即，最后这幂级数在这个区间上处处收敛到极限 0. 特别地，对于

$x=0$，级数的和必须是 0；即 $c_0=0$，因此 $a_0=b_0$. 现在我们在这个区间内部对级数微分，得到 $\phi'(x)=\sum_{\nu=1}^{\infty}\nu c_\nu x^{\nu-1}$. 然而，$\phi'(x)$ 在整个区间里也是 0；因此，特别地，对于 $x=0$，我们有 $c_1=0$ 或 $a_1=b_1$. 继续进行微分然后令 $x=0$，我们就相继求得所有的系数 c_ν 都等于零，这就证明了定理.

此外，从我们的讨论里我们能够引出下面的结论：如果我们取级数 $f(x)=\Sigma a_\nu x^\nu$ 的 ν 阶导数再令 $x=0$，我们立即得到

$$a_\nu=\frac{1}{\nu!}f^{(\nu)}(0),$$

也就是说：

每一个在除 $x=0$ 以外的点上收敛的幂级数是它所表示的函数的泰勒级数.

展开式的唯一性相当于展开式的系数能够由函数本身来表示.

*e. 解析函数

对于能够由幂级数表示的函数 $f(x)$，自从拉格朗日首先认识了这种函数的重要性以后，就被冠以“解析函数”的名字. 特别地，如果在 $x=a$ 的一个邻域里，$f(x)$ 可以展开成 $x-a$ 的收敛幂级数，$f(x)$ 就称为在 $x=a$ 的邻域里是解析的.

虽则处处不解析或并不处处解析的函数确实在分析和应用上起着很大的作用(参看第八章)，然而解析函数仍是特别重要的，因为它们与多项式一样具有许多简单的特性.

例如，一个不恒等于零的解析函数在 $x=a$ 处将有某个非零的导数. 设 r 是使 $f^{(r)}(a)\neq 0$ 的最小的数. 则 f 在点 $x=a$ 有 r 阶零点，能够表示成乘积 $f(x)=(x-a)^r g(x)$，其中 $g(x)$ 是一个解析函数，对于它 $g(a)=\frac{1}{r!}f^{(r)}(a)$ 不等于零. (与第五章第 487 页比较.)确实地，分解出幂 $(x-a)^r$ 因子的可能性从各个幂级数的收敛性可直接得到.

从 $g(x)$ 的收敛幂级数的连续性也可以看出，除非 f 恒等于零，因子 $g(x)$ 不可能在 $x=a$ 的一个适当小的邻域里为零，或者：$f(x)$ 的零点是孤

立的.

由于函数 $f'(x)$ 具有同样的性质,由此得到在一个有穷区间里一个解析函数是分段单调的,即,它不能无穷多次改变它的单调的特征;因此 $y=f(x)$ 的图形在一个有穷区间里不可能与一条直线 $y=$ 常数(或任意的直线)相交无穷多次.

可以注意的一点是,上述这些论述对于非解析函数不一定是对的,例如对于 $y=\sin(1/x)e^{-1/x^2}$, 在 $x=0$ 的邻域里(参看第 486 页),上述结论就不对.

7.6 给定函数的幂级数展开式. 待定系数法. 例

每一个幂级数在它的收敛区间内部表示一个具有各阶连续导数的连续函数. 现在我们将讨论一个给定的函数的幂级数展开的反问题. 在理论上,我们总是能够根据泰勒定理做这件事;在实践上,在 n 阶导数的实际计算和余项估计中我们经常遇到困难. 但是利用下面的对策我们经常能更简单地达到目的. 我们先试验性地写下 $f(x)=\sum_{\nu=0}^{\infty}c_\nu x^\nu$, 其中系数 c_ν 在开始时是未知的. 然后根据函数 $f(x)$ 的某些已知的性质,我们来确定这些系数,并证明该级数的收敛性. 这个级数表示一个函数,剩下只要证明这个函数是恒等于 $f(x)$ 的. 由于幂级数展开式的唯一性,我们知道除了刚才找出的那个级数以外,别的级数都不可能是所要求的展开式. 实际上,用与这一章的概念有关的一种方法,我们早已得到了 arc-tan$x$ 和 $\log(1+x)$ 的级数. 因为我们只需逐项积分这些函数的导数的级数就行了,这些级数我们已经知道它们是几何级数. 现在我们将考察关于这个方法的一些例.

a. 指数函数

正如我们在第 237 页上第三章第 4a 节看到的,函数 $y=e^x$ 的特征性质完全被微分方程 $y'=y$ 和初始条件当 $x=0$ 时 $y=1$ 决定了. 我们能够利用这些性质直接求出指数函数的幂级数. 我们的问题是要找一个函数 $f(x)$ 使得 $f'(x)=f(x)$ 和 $f(0)=1$. 如果我们试把级数写成待定系数的形式

$$f(x) = c_0 + c_1 x + c_2 x^2 + \cdots,$$

并且求导，就得到

$$f'(x) = c_1 + 2c_2 x + 3c_3 x^2 + \cdots.$$

由于根据假设这两个幂级数必须恒等，我们有等式

$$nc_n = c_{n-1},$$

对所有 $n \geq 1$ 的值都成立. 如果我们注意到由于关系式 $f(0) = 1$, 系数 c_0 必有值 1, 我们就能相继计算出所有的系数，并且得到幂级数

$$f(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots.$$

根据比值判别法，我们容易看到，这个级数对所有的 x 值收敛，因此它表示一个函数，这个函数实际上满足关系式 $f'(x) = f(x)$, $f(0) = 1$.（这里我们有意迴避了利用我们前面已经学过的有关指数函数的展开式.）

因为只有函数 e^x 具有这些性质，我们立即推知函数 $f(x)$ 恒等于 e^x.

b. 二项式级数

现在我们能够转到二项式级数了（第 479 页第 5.5c 节），这次利用待定系数法. 我们希望把函数 $f(x) = (1 + x)^\alpha$ 展开成一个幂级数，为此写下

$$f(x) = (1 + x)^\alpha = c_0 + c_1 x + c_2 x^2 + \cdots,$$

系数 c_ν 是待定的. 现在我们注意到，我们的函数明显地满足关系式

$$(1 + x)f'(x) = \alpha f(x) = \sum_{\nu=0}^{\infty} \alpha c_\nu x^\nu.$$

另一方面，如果我们对 $f(x)$ 的级数逐项微分并且乘以 $(1 + x)$, 我们就得到

$$(1 + x)f'(x) = c_1 + (2c_2 + c_1)x + (3c_3 + 2c_2)x^2 + \cdots;$$

由于 $(1 + x)f'(x)$ 的这两个幂级数必须恒等，就得到

$$\alpha c_0 = c_1, \quad \alpha c_1 = 2c_2 + c_1, \quad \alpha c_2 = 3c_3 + 2c_2, \cdots$$

现在，因为在 $x = 0$ 时我们的级数必须有值 1，我们一定有 $c_0 = 1$, 所以我们依次得到表达式

$$c_1 = 2, \quad c_2 = \frac{(\alpha - 1)\alpha}{2}, \quad c_3 = \frac{(\alpha - 2)(\alpha - 1)\alpha}{3 \cdot 2}, \cdots,$$

对于系数一般的关系式也是容易建立起来的，我们有

$$c_\nu = \frac{(\alpha - \nu + 1)(\alpha - \nu + 2)\cdots(\alpha - 1)\alpha}{\nu(\nu - 1)\cdots 2 \cdot 1} = \binom{\alpha}{\nu}.$$

将这些值代替系数，我们有级数 $\sum_{\nu=0}^{\infty}\binom{\alpha}{\nu}x^{\nu}$；不过我们还必须研究这个级数的收敛性，并且说明它确实代表 $(1+x)^{\alpha}$.

根据比值判别法我们发现，当 α 不是正整数时，这个级数当 $|x|<1$ 时收敛，当 $|x|>1$ 时发散；因为这时它的第 $n+1$ 项与第 n 项的比值为 $\frac{\alpha-n+1}{n}x$，这个表达式的绝对值当 n 无限增大时趋于 $|x|$.[1]因此，如果 $|x|<1$，我们的级数就代表一个函数 $f(x)$，根据形成系数的方法这个函数满足条件 $(1+x)f'(x)=\alpha f(x)$. 此外还有 $f(0)=1$. 这两个条件合在一起保证了函数 $f(x)$ 恒等于 $(1+x)^{\alpha}$. 因为当令

$$\phi(x)=\frac{f(x)}{(1+x)^{\alpha}},$$

我们就发现

$$\phi'(x)=\frac{(1+x)^{\alpha}f'(x)-\alpha(1+x)^{\alpha-1}f(x)}{(1+x)^{2\alpha}}=0;$$

因此 $\phi(x)$ 是一个常数，并且事实上，由于 $\phi(0)=1$，它总是等于 1 的. 于是我们证明了，对于 $|x|<1$

$$(1+x)^{\alpha}=\sum_{\nu=0}^{\infty}\binom{\alpha}{\nu}x^{\nu},$$

这就是二项式级数.

这里我们注意下面几个二项式级数的特殊情形：几何级数

$$\frac{1}{1+x}=(1+x)^{-1}=1-x+x^2-x^3+x^4-+\cdots$$

$$=\sum_{\nu=0}^{\infty}(-1)^{\nu}x^{\nu};$$

级数

$$\frac{1}{(1+x)^2}=(1+x)^{-2}=1-2x+3x^2-4x^3+-\cdots$$

---

1) 这里我们仅叙述而不证明使这个级数收敛的确切的条件. 如果指数 α 是一个 $\geqslant 0$ 的整数，级数在有限项之后结束，因此对所有的 x 值是有效的(成为通常的二项式定理). 对于 α 的所有其它的值，级数当 $|x|<1$ 时绝对收敛，当 $|x|>1$ 时发散. 当 $x=+1$ 时，如果 $\alpha>0$，级数绝对收敛；如果 $-1<\alpha<0$，级数条件收敛；如果 $\alpha\leqslant-1$，级数发散. 最后，在 $x=-1$ 处，如果 $\alpha>0$，级数绝对收敛；如果 $\alpha<0$，级数发散.

$$= \sum_{\nu=0}^{\infty} (-1)^{\nu}(\nu + 1)x^{\nu},$$

它也可以从微分几何级数得到；以及级数

$$\sqrt{(1 + x)} = (1 + x)^{1/2} = 1 + \frac{1}{2}x - \frac{1}{2 \cdot 4}x^2 + \frac{1 \cdot 3}{2 \cdot 4 \cdot 6}x^3 - \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6 \cdot 8}x^4 + - \cdots,$$

$$\frac{1}{\sqrt{(1 + x)}} = (1 + x)^{-1/2} = 1 - \frac{1}{2}x + \frac{1 \cdot 3}{2.4}x^2 - \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}x^3 + \frac{1 \cdot 3 \cdot 5 \cdot 7}{2 \cdot 4 \cdot 6 \cdot 8}x^4 - + \cdots,$$

这些级数的头两项或三项形成有用的近似式.

c. arc sin x 的级数

这个级数能够很容易地得到，方法是根据二项式级数来展开表达式 $1/\sqrt{(1 - t^2)}$：

$$(1 - t^2)^{-1/2} = 1 + \frac{1}{2}t^2 + \frac{1 \cdot 3}{2 \cdot 4}t^4 + \cdots.$$

当 $|t| < 1$ 时这个级数收敛，因此当 $|t| \leqslant q < 1$ 时一致收敛. 在 0 到 x 之间逐项积分，我们得到

$$\arcsin x = x + \frac{1}{2}\frac{x^3}{3} + \frac{1 \cdot 3}{2 \cdot 4}\frac{x^5}{5} + \cdots;$$

根据比值判别法我们知道当 $|x| < 1$ 时级数收敛，当 $|x| > 1$ 时级数发散.

若是从泰勒定理来推导这个级数就会很不方便，因为余项估计很困难.

d. ar sinh $x = \log[x + \sqrt{(1 + x^2)}]$ 的级数

我们用类似的方法求得这个展开式. 利用二项式定理我们写下 ar sinh x 的导数的级数，

$$\frac{1}{\sqrt{1 + x^2}} = 1 - \frac{1}{2}x^2 + \frac{1 \cdot 3}{2 \cdot 4}x^4 - \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}x^6 + - \cdots,$$

然后逐项积分. 这样我们就得到展开式

$$\text{ar sinh}\, x = x - \frac{1}{2}\frac{x^3}{3} + \frac{1 \cdot 3}{2 \cdot 4}\frac{x^5}{5} - + \cdots,$$

它的收敛区间是 $-1 \leqslant x \leqslant 1$.

e. 级数乘法的例

函数

$$\frac{\log(1+x)}{1+x}$$

的展开式是应用幂级数乘法规则的一个简单例子. 我们只需将对数级数

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + - \cdots$$

乘以几何级数

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 - + \cdots;$$

正如读者自己可以证实的那样，对于 $|x| < 1$，我们得到值得注意的展开式

$$\frac{\log(1+x)}{1+x} = x - \left(1 + \frac{1}{2}\right)x^2 + \left(1 + \frac{1}{2} + \frac{1}{3}\right)x^3$$

$$- \left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4}\right)x^4 + - \cdots.$$

f. 逐项积分的例(椭圆积分)

在前面第 319 和 434 页的应用中,我们已经遇见过椭圆积分

$$K = \int_0^{\frac{\pi}{2}} \frac{d\phi}{\sqrt{(1 - k^2\sin^2\phi)}} \qquad (k^2 < 1)$$

[一个摆的振动周期]. 为了计算这个积分,我们可以先按二项式定理展开被积函数,得到

$$\frac{1}{\sqrt{(1 - k^2\sin^2\phi)}} = 1 + \frac{1}{2}k^2\sin^2\phi + \frac{1\cdot 3}{2\cdot 4}k^4\sin^4\phi$$

$$+ \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}k^6\sin^6\phi + \cdots.$$

由于 $k^2\sin^2\phi$ 决不会大于 k^2，因此这个级数对于所有的 ϕ 值一致收敛，因而我们可以逐项积分:

$$K = \int_0^{\frac{\pi}{2}} \frac{d\phi}{\sqrt{(1 - k^2\sin^2\phi)}} = \int_0^{\frac{\pi}{2}} d\phi + \frac{1}{2}k^2\int_0^{\frac{\pi}{2}} \sin^2\phi\, d\phi$$

$$+\frac{1\cdot 3}{2\cdot 4}k^4\int_0^{\frac{\pi}{2}}\sin^4\phi d\phi+\cdots.$$

这里出现的积分早已计算过了[参看第 298 页等式(76)]. 代入它们的值,我们就得到

$$K=\int_0^{\frac{\pi}{2}}\frac{d\phi}{\sqrt{(1-k^2\sin^2\phi)}}$$

$$=\frac{\pi}{2}\left[1+\left(\frac{1}{2}\right)^2k^2+\left(\frac{1\cdot 3}{2\cdot 4}\right)^2k^4+\left(\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\right)^2k^6+\cdots\right].$$

7.7 复数项幂级数

a. 在幂级数中引进复数项. 三角函数的复数表示式

有一些在表面上没有什么关系的函数，其幂级数之间的相似性引导欧拉得到了这些函数之间的纯形式的联系，其办法是给变量 x 以复值，特别是给 x 以纯虚值. 我们先叙述欧拉形式的但引人注目的且富有成果的发现，而不受严格性的妨碍，然后我们再指出一个更严格的验证.

如果在 e^x 的级数里我们用一个纯虚数 $i\phi$，其中 ϕ 是一个实数，代替量 x，就得到这种类型的第一个关系式. 如果我们回忆起虚单位 i 的基本方程，即，$i^2=-1$，从而 $i^3=-i$，$i^4=1$，$i^5=i$，$\cdots$ 等等，然后将级数的实项和虚项分开，我们就得到

$$e^{i\phi}=\left(1-\frac{\phi^2}{2!}+\frac{\phi^4}{4!}-\frac{\phi^6}{6!}+-\cdots\right)$$

$$+i\left(\phi-\frac{\phi^3}{3!}+\frac{\phi^5}{5!}-\frac{\phi^7}{7!}+-\cdots\right),$$

或者换一种方式，

$$e^{i\phi}=\cos\phi+i\sin\phi. \tag{11}$$

这就是熟知且重要的"欧拉公式"，它是分析上的一个里程碑；虽然到目前为止它还纯粹是形式的.[1]这与棣美弗定理(第 111 页)是一致的，棣美弗定理由方程式

---

1) 对于 $\phi=\pi$ 的一个结果是公式 $e^{\pi i}=-1$，这是三个最重要的常数 e、π 和 i 之间的一个明显的关系式.

$$(\cos\phi + i\sin\phi)(\cos\psi + i\sin\psi)$$
$$= \cos(\phi + \psi) + i\sin(\phi + \psi)$$
表示．根据欧拉公式，这个等式仅说明关系式
$$e^x \cdot e^y = e^{x+y}$$
对于纯虚值 $x = i\phi$，$y = i\psi$ 仍然成立．

应当说，只需定义 $e^{i\phi}$ 为复数 $\cos\phi + i\sin\phi$，这个欧拉公式和加法定理 $e^{i\phi}e^{i\psi} = e^{i(\phi+\psi)}$ 就可以严格地使用而无需进一步验证．这个定义和指数运算的通常规则是一致的．特别地，e 的乘幂的通常法则正好给出了用棣美弗公式表示的三角学的加法定理的简明表达式，而棣美弗公式本身却完全是一个初等性质的公式．因此当我们不借助更一般的复变数函数论而利用欧拉的关系式时，如同在下一节中那样，我们是建立在可靠的基础上的．

更一般地，我们能够对于任意的复指数 $x + iy$（其中 x 和 y 是实的）用公式
$$e^{x+iy} = e^x \cdot e^{iy} = e^x(\cos y + i\sin y)$$
定义指数函数．

如果我们在 $\cos x$ 的幂级数里用纯虚数 ix 代替变量 x，我们立即得到 $\cosh x$ 的级数；这个关系能够用等式
$$\cosh x = \cos ix \tag{12}$$
来表示．用同样的方法我们得到
$$\sinh x = \frac{1}{i}\sin ix. \tag{13}$$

由于欧拉公式也给出 $e^{-i\phi} = \cos\phi - i\sin\phi$，我们得到三角函数的指数表达式，
$$\sin x = \frac{e^{ix} - e^{-ix}}{2i}, \qquad \cos x = \frac{e^{ix} + e^{-ix}}{2}. \tag{14}$$
这些恰好类似于双曲函数的指数表达式，而且事实上，它们通过关系式 $\cosh x = \cos ix$，$\sinh x = \frac{1}{i}\sin ix$ 就变换成双曲函数．

当然，对于函数 $\tan x$，$\tanh x$，$\cot x$，$\coth x$ 也能得到相应的

形式关系，它们是由等式 $\tanh x = \frac{1}{i}\tan ix$，$\coth x = i\cot ix$ 联系起来的。

最后，对于反三角函数和反双曲函数也能得到类似的关系。例如，从

$$y = \tan x = \frac{e^{ix} - e^{-ix}}{i(e^{ix} + e^{-ix})} = \frac{e^{2ix} - 1}{i(e^{2ix} + 1)}$$

我们立即求得

$$e^{2ix} = \frac{1 + iy}{1 - iy}.$$

如果我们在这个等式的两边同时取对数，然后用 x 代替 y 和用 $\arctan x$ 代替 x，我们就得到等式

$$\arctan x = \frac{1}{2i}\log\frac{1 + ix}{1 - ix}, \tag{15}$$

这个等式表示了反正切和对数之间的一个值得注意的关系。如果在关于 $\frac{1}{2}\log\frac{1+x}{1-x}$ 的熟知的幂级数里(第 467 页)，用 ix 代替 x，我们实际上就得到了关于 $\arctan x$ 的幂级数，

$$\arctan x = \frac{1}{i}\left(ix + \frac{(ix)^3}{3} + \frac{(ix)^5}{5} + \cdots\right)$$

$$= x - \frac{x^3}{3} + \frac{x^5}{5} - + \cdots.$$

这些关系到目前为止仍不过是属于纯粹形式性质的，因而自然地需要对于企图表达的意义有一个更严格的叙述。然而，我们在上面已经看到，利用适当的定义，这些关系就获得了一个满意的严格的意义。

*b. 复变函数一般理论一瞥

虽然在上一节指出的纯形式的观点本身是无可非议的，但是人们仍期望在前面的公式里认识一些比仅仅是形式联系更多的东西。这个目标引来了**复函数的一般理论**，我们(为了简洁起见)称

之为所谓一个复变量的解析函数的一般理论．我们可以用具有复变量和复系数的幂级数理论的一般讨论作为我们的出发点．一旦我们在复数范围里定义了极限的概念，这样一个幂级数的理论的构造是并不困难的；事实上，它几乎严格地平行于实的幂级数的理论．然而，因为在下文中我们将不使用这些内容的任何部分，所以在这里我们只局限于叙述某些事实而省略证明．人们发现 7.5a 节的定理的下述推广对于复的幂级数是成立的：

如果一个幂级数对任一复数值 $x=\xi$ 收敛，那末对于每一个满足 $|x|<|\xi|$ 的 x 值这个级数绝对收敛；如果对于一个值 $x=\xi$ 级数发散，那末对于每一个满足 $|x|>|\xi|$ 的 x 值这个级数发散．一个幂级数不处处收敛，但是在除 $x=0$ 以外的某一点处收敛，则存在一个收敛圆，即，存在一个数 $\rho>0$ 使得级数在 $|x|<\rho$ 时绝对收敛而在 $|x|>\rho$ 时发散．

一旦建立起了用幂级数表示的复变量的函数概念，并且发展了对于这样的函数的运算法则，我们就能够认为，复的变量 x 的函数 e^x，$\sin x$，$\cos x$，$\arctan x$ 等等就是用当 x 取实值时表示这些函数的幂级数来简单地定义的．

我们将举两个例来指出，复变量的这一引进是怎样说明初等函数的性质的．对于 $\frac{1}{1+x^2}$，它的几何级数当 x 离开区间 $-1\leqslant x\leqslant 1$ 时是不收敛的，从而 $\arctan x$ 的级数同样也是如此，虽然这些函数在收敛区间的端点的性质没有什么特殊；事实上，它们和它们的各阶导数对于所有实的 x 值都是连续的．另一方面，我们能够容易地理解，对于 $\frac{1}{1-x^2}$ 和 $\log(1-x)$，它们的级数当 x 通过值 1 的时候是不收敛的，因为在那里它们成了无穷．

但是如果我们也考虑 x 的复数值，则对于反正切的级数和级数 $\sum_{\nu=0}^{\infty}(-1)^{\nu}x^{2\nu}$ 当 $|x|>1$ 时的发散性立刻就成为明显的了．因为我们发现，当 $x=i$ 时，这些函数成为无穷，所以不能够用一个收敛级数来表示．因此根据我们的关于收敛圆的定理，这些级数

对于所有使得 $|x| > |i| = 1$ 的 x 值一定发散;特别是,对于 x 的实数值级数,在区间 $-1 \leqslant x \leqslant 1$ 之外发散.

另一个例是已研究过的函数 $f(x) = e^{-1/x^2}(x \neq 0)$, $f(0) = 0$(参看第 486 页),这个函数尽管它是完全光滑的,但不能展开成泰勒级数. 事实上,如果我们把纯虚值 $x = i\xi$ 考虑进来,这个函数就不连续了. 这时这函数具有形式 e^{1/ξ^2},并且当 $\xi \to 0$ 时无限增大. 因此显然,在原点的邻域里,无论我们选择怎样小的一个邻域,对所有的复的 x 值,没有一个 x 的幂级数能够代表这个函数.

在函数理论上和复变量的幂级数方面的这些评注,在这里已足以满足我们的需要了.

附　　录

*A.1 级数的乘法和除法

a. 绝对收敛级数的乘法

设
$$A = \sum_{\nu=0}^{\infty} a_\nu, \qquad B = \sum_{\nu=0}^{\infty} b_\nu$$
是两个绝对收敛的级数. 与这些级数一起我们考虑对应的绝对值的收敛级数
$$\bar{A} = \sum_{\nu=0}^{\infty} |a_\nu| \quad 和 \quad \bar{B} = \sum_{\nu=0}^{\infty} |b_\nu|.$$
我们进一步设
$$A_n = \sum_{\nu=0}^{n-1} a_\nu, \quad B_n = \sum_{\nu=0}^{n-1} b_\nu, \quad \bar{A}_n = \sum_{\nu=0}^{n-1} |a_\nu|, \quad \bar{B}_n = \sum_{\nu=0}^{n-1} |b_\nu|,$$
以及
$$c_n = a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0.$$
我们断言,级数 $\sum_{\nu=0}^{\infty} c_\nu$ 是绝对收敛的,并且它的和等于 AB.

为了证明这点,我们写下级数
$$a_0b_0 + a_1b_0 + a_1b_1 + a_0b_1 + a_2b_0 + a_2b_1 + a_2b_2 + a_1b_2 + a_0b_2 + \cdots$$
$$+ a_nb_0 + a_nb_1 + \cdots + a_nb_n + \cdots + a_1b_n + a_0b_n + \cdots,$$
这个级数的第 n^2 个部分和是 A_nB_n; 我们断言,这个级数绝对收敛. 因为对应的绝对值的级数的部分和单调递增;它的第 n^2 个部分和等于 $\bar{A}_n\bar{B}_n$,它小

于 $\bar{A}\bar{B}$（并且它趋向 $\bar{A}\bar{B}$）. 因此绝对值的级数收敛，因而上面写下的级数绝对收敛. 级数的和显然是 AB，因为它的第 n^2 个部分和是 A_nB_n，当 $n\to\infty$ 时，A_nB_n 趋于 AB. 我们现在交换项的次序，这对于绝对收敛的级数是允许的；我们还把接连的一些项括在一起. 在一个收敛级数里，我们可以在任意多的地方将接连的一些项括在一起而不影响级数的收敛性或改变级数的和，因为如果我们括在一起的所有项，譬如是 $(a_{n+1}+a_{n+2}+\cdots+a_m)$，那末当我们形成部分和时我们将省去原来落在 s_n 和 s_m 之间的那些部分和，这并不影响收敛性或改变极限的值. 此外，如果级数在加进括号之前是绝对收敛的，那末加括号后它仍然绝对收敛. 因为级数

$$\sum_{\nu=0}^{\infty} c_\nu = (a_0b_0) + (a_0b_1 + a_1b_0) + (a_0b_2 + a_1b_1 + a_2b_0) + \cdots$$

是由上面写下的级数按这种方式形成的，因此所需要的证明完成了.

*b. 幂级数的乘法和除法

我们的定理主要应用在幂级数的理论上. 下面的论断是它的一个直接推论：在两个幂级数

$$\sum_{\nu=0}^{\infty} a_\nu x^\nu \quad 和 \quad \sum_{\nu=0}^{\infty} b_\nu x^\nu$$

的公共收敛区间里，这两个幂级数的乘积可以由第三个幂级数 $\sum\limits_{\nu=0}^{\infty} c_\nu x^\nu$ 表示，此幂级数的系数由

$$c_\nu = a_0b_\nu + a_1b_{\nu-1} + \cdots + a_\nu b_0$$

给出.

*至于幂级数的除法，我们也同样能够用一个幂级数 $\sum\limits_{\nu=0}^{\infty} q_\nu x^\nu$ 来表示上面的两个幂级数的商，不过要分母的常数项 b_0 不为零.（在 b_0 为零的情况下，这样的一种表示法一般说来是不可能的；因为在 $x=0$ 时，由于分母为零，级数不可能收敛，而另一方面，每一个幂级数却必须在 $x=0$ 时收敛.）幂级数

$$\sum_{\nu=0}^{\infty} q_\nu x^\nu$$

的系数能够根据 $\sum\limits_{\nu=0}^{\infty} q_\nu x^\nu \cdot \sum\limits_{\nu=0}^{\infty} b_\nu x^\nu = \sum\limits_{\nu=0}^{\infty} a_\nu x^\nu$ 来计算，因此下面的方程必须成立：

$$
\begin{aligned}
a_0 &= q_0 b_0, \\
a_1 &= q_0 b_1 + q_1 b_0, \\
a_2 &= q_0 b_2 + q_1 b_1 + q_2 b_0, \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
a_\nu &= q_0 b_\nu + q_1 b_{\nu-1} + \cdots + q_\nu b_0.
\end{aligned}
$$

从这些方程的第一个，容易求得 q_0 的值，从第二个方程我们求得 q_1 的值，从第三个方程（利用 q_0 和 q_1 的值）我们求得 q_2 的值，等等. 为了给出由第三个幂级数表示的两个幂级数商的表达式的严格证明，我们必须研究形式地计算出来的幂级数 $\sum_{\nu=0}^{\infty} q_\nu x^\nu$ 的收敛性. 但是，我们将不再进一步使用这个结果，因而我们就满足于陈述商的级数在原点的某个区间内确实收敛，证明就不讲了.

A.2 无穷级数与广义积分

无穷级数以及和它们相联系而发展出来的概念在广义积分理论上有简单的应用和类比(参看第三章第 3.15节). 我们局限在具有无穷积分区间的收敛积分的情形，譬如说形如 $\int_0^\infty f(x)\,dx$ 的积分. 如果我们用一个单调趋于 $+\infty$ 的数列 $x_0 = 0$, $x_1, \cdots$ 去分割积分区间，我们就能够将这广义积分写成这样:

$$\int_0^\infty f(x)\,dx = a_1 + a_2 + \cdots,$$

这里我们的无穷级数的每一项都是一个积分;

$$a_1 = \int_0^{x_1} f(x)\,dx, \qquad a_2 = \int_{x_1}^{x_2} f(x)\,dx, \cdots$$

等等. 无论我们选择怎样的点 x_ν 这都是对的. 因此我们能够用许多种方法将收敛的广义积分的概念与一个无穷级数的概念联系起来.

特别方便的是选择点 x_ν 使得积分在每一个子区间里都不改变符号. 于是级数 $\sum_{\nu=1}^{\infty} |a_\nu|$ 就对应于我们的函数的绝对值的积分

$$\int_0^\infty |f(x)|\,dx.$$

这样我们就自然地引出下面的概念：如果积分 $\int_0^\infty |f(x)|\,dx$ 收敛，则广义积分 $\int_0^\infty f(x)\,dx$ 就称为是绝对收敛的。否则，如果我们的积分存在，我们就说这广义积分是条件收敛的。

早已研究过的一些积分（第 329 页到第 333 页），例如

$$\int_0^\infty \frac{1}{1+x^2}\,dx\,,\quad \int_0^\infty e^{-x^2}\,dx\,,\quad \Gamma(x)=\int_0^\infty e^{-t}t^{x-1}\,dx\,,$$

是绝对收敛的。

另一方面，在第 333 页中研究过的重要的"狄里克莱"积分

$$J=\int_0^\infty \frac{\sin x}{x}\,dx=\lim_{A\to\infty}\int_0^A \frac{\sin x}{x}\,dx\,,$$

是一个条件收敛积分的典型例子。收敛性的最简单的证明是化成一个绝对收敛的积分：我们写成 $\sin x=(1-\cos x)'=2\left(\sin^2\frac{x}{2}\right)'$，并且利用分部积分法，将 J 变成绝对收敛的形式

$$J=2\int_0^\infty\left(\sin^2\frac{x}{2}\right)\frac{1}{x^2}\,dx\,.$$

（注意，当 $x\to 0$ 时，新的被积函数连续地通过极限值 $\frac{1}{2}$，并且当 $x\to\infty$ 时按 x^{-2} 阶消失为零。）

*收敛性的另一个证明是用点 $x_\nu=\nu\pi\ (\nu=0,\ 1,\ 2,\cdots,\ \mu_A)$ 分割从 0 到 A 的区间，其中 μ_A 是满足 $\mu_A\pi\leqslant A$ 的最大的整数。于是积分被分成形如

$$a_\nu=\int_{(\nu-1)\pi}^{\nu\pi}\frac{\sin x}{x}\,dx\qquad(\nu=1,\ 2,\ \cdots)$$

的项和形如

$$\int_{\mu_A\pi}^{A}\frac{\sin x}{x}\,dx\qquad(0\leqslant A-\mu_A\pi<\pi)$$

的一个余项 R_A。

明显地，量 a_ν 交替地取正负，因为 $\sin x$ 在相继的区间里交替地取正和负。而且，$|a_{\nu+1}|<|a_\nu|$；因为运用变换 $x=\xi-\pi$，

就有

$$|a_\nu| = \int_{(\nu-1)\pi}^{\nu\pi} \frac{|\sin x|}{x} dx = \int_{\nu\pi}^{(\nu+1)\pi} \frac{|\sin(\xi-\pi)|}{\xi-\pi} d\xi$$

$$= \int_{\nu\pi}^{(\nu+1)\pi} \frac{|\sin \xi|}{\xi-\pi} d\xi > \int_{\nu\pi}^{(\nu+1)\pi} \frac{|\sin \xi|}{\xi} d\xi = |a_{\nu+1}|.$$

因此，根据莱布尼兹判别法，我们知道 Σa_ν 收敛．而且，余项 R_A 有绝对值

$$|R_A| = \left|\int_{\mu_A\pi}^{A} \frac{\sin x}{x} dx\right| \leqslant \int_{\mu_A\pi}^{(\mu_A+1)\pi} \frac{|\sin x|}{x} dx$$

$$\leqslant \frac{1}{\mu_A\pi} \int_{\mu_A\pi}^{(\mu_A+1)\pi} |\sin x|\, dx = \frac{2}{\mu_A\pi},$$

而当 A 增大时这是趋于 0 的．因此，如果我们在等式

$$\int_0^A \frac{\sin x}{x} dx = a_1 + a_2 + a_3 + \cdots + a_{\mu_A} + R_A$$

中让 A 趋于 ∞，右边就趋于 Σa_ν．并以它作为极限，因而我们的积分是收敛的．但这收敛性不是绝对的，因为

$$|a_\nu| > \int_{(\nu-1)\pi}^{\nu\pi} \frac{|\sin x|}{\nu\pi} dx = \frac{2}{\nu\pi},$$

因此 $\Sigma|a_\nu|$ 发散．

*A.3 无穷乘积

在这一章的引言中(第 536 页)，我们谈到过无穷级数虽然是一种特别重要的方法，只是用无穷过程表示数或函数的**一种**方法．作为这种过程的另一个例子，我们考虑无穷乘积．不再给出证明．

在第 298 页我们遇见过瓦里斯乘积

$$\frac{\pi}{2} = \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdots,$$

在这个式子里，数 $\frac{\pi}{2}$ 表示成了一个“无穷乘积”．一般地讲，所谓

无穷乘积

$$\prod_{\nu=1}^{\infty} a_\nu = a_1 \cdot a_2 \cdot a_3 \cdot a_4 \cdots$$

的值，只要它的极限存在，我们指的就是“部分乘积”的序列

$$a_1, \quad a_1 \cdot a_2, \quad a_1 \cdot a_2 \cdot a_3, \quad a_1 \cdot a_2 \cdot a_3 \cdot a_4, \cdots$$

的极限.

当然，因子 $a_1, a_2, a_3, \cdots$ 也可以是一个变量 x 的函数. 一个特别有趣的例是关于函数 $\sin x$ 的“无穷乘积”,

$$\sin \pi x = \pi x \left(1 - \frac{x^2}{1^2}\right)\left(1 - \frac{x^2}{2^2}\right)\left(1 - \frac{x^2}{3^2}\right)\cdots, \qquad (16)$$

这个式子我们将在第 635 页 8.5 节得到.

在数论中关于 zeta 函数的无穷乘积起着非常重要的作用. 为了沿用在数论中常用的记号，这里我们用 s 表示自变量，并且按照黎曼意义对于 $s > 1$ 我们用表达式

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

来定义 zeta 函数. 我们知道(第 550 页 7.2c 节)，如果 $s > 1$，右边的级数收敛. 如果 p 是任一个大于 1 的数，我们得到等式

$$\frac{1}{1 - \frac{1}{p^s}} = 1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \frac{1}{p^{3s}} + \cdots.$$

这是把左边展开成一个公比为 p^{-s} 的几何级数. 如果我们设想这个级数对所有的素数 $p_1, p_2, p_3, \cdots$ 按它们从小到大的顺序写下来，并且所有这样形成的等式乘在一起，我们就在左边得到一个乘积，形如

$$\frac{1}{1 - p_1^{-s}} \cdot \frac{1}{1 - p_2^{-s}} \cdot \frac{1}{1 - p_3^{-s}} \cdots.$$

我们不来验证这个过程，而将等式右边的级数乘在一起，我们就得到一个和，它的各项为

$$p_1^{-k_1 s} p_2^{-k_2 s} p_3^{-k_3 s} \cdots = (p_1^{k_1} p_2^{k_2} p_3^{k_3} \cdots)^{-s},$$

其中 $k_1, k_2, k_3, \cdots$ 是任意的非负整数. 我们也还记得，根据一个基本定理，每一个整数 $n > 1$ 能且只能用一种方式表示成不同的素数的幂的乘积 $n = p_1^{k_1} p_2^{k_2} \cdots$. 这样我们就发现等式右边的乘积就是函数 $\zeta(s)$，从而我们

得到值得注意的欧拉的"乘积形式"

$$\zeta(s)=\frac{1}{1-p_1^{-s}}\cdot\frac{1}{1-p_2^{-s}}\cdot\frac{1}{1-p_3^{-s}}\cdots. \tag{17}$$

这个"乘积形式"，它的推导我们在这里只简略地作了概述. 因为素数的个数是无穷的，所以它的确是 zeta 函数表示成无穷乘积的一种表示法.

在无穷乘积的一般理论里经常排除乘积 $a_1a_2\cdots a_n$ 取极限零这种情形. 因此没有一个因子 a_n 是零的情形是特别重要的. 为了乘积能够收敛，当 n 增大时因子 a_n 必须趋于 1. 因为如果需要，我们可以省去有穷个因子（这与收敛性的问题无关），所以我们可以假定 $a_n>0$. 下面几乎是显然的定理适用于这种情形：

当 $a_\nu>0$ 时，乘积 $\prod_{\nu=1}^{\infty}a_\nu$ 收敛的一个必要充分条件是级数 $\sum_{\nu=1}^{\infty}\log a_\nu$ 应当收敛. 因为作为对数的连续性的一个推论，当且仅当部分乘积 $a_1a_2\cdots a_n$ 具有一个正的极限时，这个级数的部分和 $\sum_{\nu=1}^{n}\log a_\nu=\log(a_1a_2\cdots a_n)$ 才趋向一个确定的极限.

当 $a_\nu=1+\alpha_\nu$ 时，在收敛性的研究中下面的充分条件经常适用. 如果级数

$$\sum_{\nu=1}^{\infty}|\alpha_\nu|$$

收敛并且没有一个因子 $(1+\alpha_\nu)$为零，则乘积

$$\prod_{\nu=1}^{\infty}(1+\alpha_\nu)$$

收敛. 在证明中，我们可以假设每一个 $|\alpha_\nu|<\frac{1}{2}$，因为如果必要可略去有限个因子. 于是我们有 $1-|\alpha_\nu|>\frac{1}{2}$. 根据中值定理 $\log(1+h)=\log(1+h)-\log 1=h/(1+\theta h)$，其中 $0<\theta<1$，于是

$$|\log(1+\alpha_\nu)| = \left|\frac{\alpha_\nu}{1+\theta\alpha_\nu}\right| \leqslant \frac{|\alpha_\nu|}{1-|\alpha_\nu|} \leqslant 2|\alpha_\nu|,$$

从而由 $\sum_{\nu=1}^{\infty}|\alpha_\nu|$ 的收敛性得出级数 $\sum_{\nu=1}^{\infty}\log(1+\alpha_\nu)$ 的收敛性。

从我们的判别准则可以推知，上面关于 $\sin\pi x$ 的无穷乘积(16)对于所有的 x 值是收敛的，例外的只在于 $x=0,\ \pm 1,\ \pm 2,\ \cdots$，这时乘积的因子是零. 至于黎曼 ζ 函数，对于 $p \geqslant 2$ 和 $s>1$，我们容易求得

$$\frac{1}{1-p^{-s}} = 1+\frac{1}{p^s-1}, \qquad 0<\frac{1}{p^s-1}<\frac{2}{p^s}.$$

现在如果我们让 p 取一切素数值，级数 $\Sigma\frac{1}{p^s}$ 一定收敛，因为它的项只形成了收敛级数 $\sum_{\nu=1}^{\infty}\frac{1}{\nu^s}$ 的一部分. 于是在方程(17)中的乘积对于 $s>1$ 的收敛性得证. 根据当 $s=1$ 的 $\zeta(s)$ 的级数(即调和级数)发散这一事实，我们能够引出一个值得注意的结论：素数的倒数的级数，即级数

$$\sum_{k=1}^{\infty}\frac{1}{p_k} = \frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\frac{1}{7}+\frac{1}{11}+\frac{1}{13}+\frac{1}{17}+\frac{1}{19}+\cdots$$

发散. (顺便说说，这表明了素数的个数是无穷的.)的确，假如素数的倒数的级数是收敛的，那末以

$$\alpha_k = \frac{1}{1-p_k^{-1}}-1 = \frac{p_k^{-1}}{1-p_k^{-1}}$$

为项的级数也应收敛，因为 $p_k \geqslant 2$ 和

$$0<\alpha_k \leqslant \alpha p_k^{-1}.$$

于是，根据我们的判别法，无穷乘积

$$\prod_{k=1}^{\infty}(1+\alpha_k) = \prod_{k=1}^{\infty}\frac{1}{1-p_k^{-1}} = \prod_{k=1}^{\infty}\left(1+\frac{1}{p_k}+\frac{1}{p_k^2}+\cdots\right)$$

也应收敛；于是明显地调和级数同样也应收敛，而这是不可能的.

*A.4 含有伯努利数的级数

到目前为止对于某些初等函数，例如 $\tan x$，我们还没有给出幂级数的展开式. 原因是它的数字系数不以任何简单的形式出现. 我们能够用所谓的**伯努利数**来表示这些系数和其他一些函数的级数里的那些系数. 这些数是稀奇的有理数，它们的形成规律有点隐蔽，它们在分析的许多部分中出现.

要达到用伯努利数表示级数的系数这个目的，最简单的方法是将函数

$$\frac{x}{e^x-1}=\frac{1}{1+\frac{x}{2!}+\frac{x^2}{3!}+\cdots}$$

展开成一个形如

$$\frac{x}{e^x-1}=\sum_{\nu=0}^{\infty}\frac{B_\nu^*}{\nu!}x^\nu$$

的幂级数．如果我们将这个等式写成

$$x=(e^x-1)\sum_{\nu=0}^{\infty}\frac{B_\nu^*}{\nu!}x^\nu$$

并且将 e^x-1 的幂级数代入等式右边，我们就得到关于 B_n^* 的递推关系

$$\binom{n+1}{1}B_n^*+\binom{n+1}{2}B_{n-1}^*+\binom{n+1}{3}B_{n-2}^*+\cdots+\binom{n+1}{n+1}B_0^*=0$$

对于 $n>0$ 成立，而 $B_0^*=1$．从这些关系式中 B_n^* 能够容易地相继计算出来． 这些有理数称为伯努利数 (Bernoulli numbers)．[1]它们之所以是有理数，是因为在它们的形成中只涉及到有理运算； 容易看到， 所有的奇数项除了 $\nu=1$ 以外都等于零．其最前面几个是

$$B_0^*=1,\quad B_1^*=-\frac{1}{2},\quad B_2^*=\frac{1}{6},\quad B_4^*=-\frac{1}{30},$$

$$B_6^*=\frac{1}{42},\quad B_8^*=-\frac{1}{30},\quad B_{10}^*=\frac{5}{66},\cdots$$

至于这些数在我们的问题里是如何出现在幂级数中的，我们只能简略讲．首先，利用变换

$$1+\frac{B_2^*}{2!}x^2+\cdots=\frac{x}{e^x-1}+\frac{x}{2}=\frac{x}{2}\cdot\frac{e^x+1}{e^x-1}$$

$$=\frac{x}{2}\cdot\frac{e^{\frac{1}{2}x}+e^{-\frac{1}{2}x}}{e^{\frac{1}{2}x}-e^{-\frac{1}{2}x}},$$

我们得到

---

1) 在稍有不同的记法里(第 659 页)，基本公式写成

$$\frac{x}{e^x-1}=1-\frac{1}{2}x+\sum_{\nu=1}^{\infty}(-1)^{\nu+1}\frac{B_\nu}{(2\nu)!}x^{2\nu}.$$

$$\frac{x}{2}\coth\frac{x}{2}=\sum_{\nu=0}^{\infty}\frac{B_{2\nu}^{*}}{(2\nu)!}x^{2\nu}.$$

(这个公式证明了当 $\nu>0$ 时 $B_{2\nu+1}^{*}=0$，因为 $\frac{x}{2}\coth\frac{x}{2}$ 是 x 的一个偶函数.)

如果我们用 $2x$ 代替 x，我们就有级数

$$x\coth x=\sum_{\nu=0}^{\infty}\frac{2^{2\nu}B_{2\nu}^{*}}{(2\nu)!}x^{2\nu},$$

可以证明，这对于 $|x|<\pi$ 成立；从这个等式里，用 $-ix$ 代替 x，我们得到（参看第 580 页）

$$x\cot x=\sum_{\nu=0}^{\infty}(-1)^{\nu}\frac{2^{2\nu}B_{2\nu}^{*}}{(2\nu)!}x^{2\nu},\qquad |x|<\pi.$$

现在，根据等式 $2\cot 2x=\cot x-\tan x$，我们得到级数

$$\tan x=\sum_{\nu=1}^{\infty}(-1)^{\nu-1}\frac{2^{2\nu}(2^{2\nu}-1)}{(2\nu)!}B_{2\nu}^{*}x^{2\nu-1},$$

它对于 $|x|<\frac{\pi}{2}$ 成立.

关于更进一步的知识，我们建议读者参考第八章和更详细的著作.[1]

问　　题

7.1 节，第 537 页

1．证明

$$\sum_{\nu=1}^{\infty}\frac{1}{\nu(\nu+1)}=\frac{1}{1.2}+\frac{1}{2.3}+\cdots=1.$$

[参看 1.6 节的问题 12(a)] 并且利用此结果证明 $\sum_{\nu=1}^{\infty}\frac{1}{\nu^{2}}$ 收敛.

2．利用第一题的结果得出

$$\sum_{\nu=1}^{\infty}\frac{1}{\nu^{2}}$$

1）例如参看 K. Knopp 著《无穷级数的理论和应用》，第 183 页，Blackie & Son 1928 年出版，和 K. Knopp 著《无穷序列和无穷级数》，1956 年 Dover 发行公司.

的上界和下界.

3. 证明 $\sum_{\nu=0}^{\infty}(-1)^{\nu}\frac{2\nu+3}{(\nu+1)(\nu+2)}=1$.

4. 什么样的 α 值使级数 $1-\frac{1}{2^{\alpha}}+\frac{1}{3^{\alpha}}-\frac{1}{4^{\alpha}}+-\cdots$ 收敛?

5. 证明如果 $\sum_{\nu=1}^{\infty}a_{\nu}$ 收敛,并且 $s_n=a_1+a_2+\cdots+a_n$,则序列

$$\frac{s_1+s_2+\cdots+s_N}{N}$$

也收敛,并且以 $\sum_{\nu=1}^{\infty}a_{\nu}$ 作为它的极限.

6. 级数 $\sum_{n=1}^{\infty}\left(\frac{2n}{2n+1}-\frac{2n-1}{2n}\right)$ 是否收敛?

7. 级数 $\sum_{\nu=1}^{\infty}(-1)^{\nu}\frac{\nu}{\nu+1}$ 是否收敛?

8. 证明如果 $\sum_{\nu=1}^{\infty}a_{\nu}^2$ 收敛,则 $\sum_{\nu=1}^{\infty}\frac{a_{\nu}}{\nu}$ 也收敛.

9. (a) 如果 a_n 是正项单调递增序列,什么时候级数

$$\frac{1}{a_1}+\frac{1}{a_1a_2}+\frac{1}{a_1a_2a_3}+\cdots \text{ 收敛?}$$

(b) 举出一个单调递降的序列具有 $\lim_{n\to\infty}a_n=1$,使得这个序列所对应的上述级数发散.

(c) 证明如果序列是递减的,那末甚至当 $\lim_{n\to\infty}a_n=1$ 时也可能在(a)中得到收敛的和.

10. 如果具有递减的正项的级数 $\sum_{\nu=1}^{\infty}a_{\nu}$ 收敛,则 $\lim_{n\to\infty}na_n=0$.

11. 证明级数 $\sum_{\nu=1}^{\infty}\sin\frac{\pi}{\nu}$ 发散.

12. 证明如果 Σa_{ν} 收敛并且 $b_1, b_2, b_3, \cdots$ 是一个有界的单调数列,则 $\Sigma a_{\nu}b_{\nu}$ 收敛. 再证明如果 $s=\Sigma a_{\nu}b_{\nu}$ 并且 $\Sigma a_{\nu}\leqslant M$,则 $|s|\leqslant Mb_1$.

13. 如果级数

$$\sum_{i=1}^{\infty}|a_{i+1}-a_i|$$

收敛,就说序列 $\{a_n\}$ 有有界变差.

(a) 证明如果序列 $\{a_n\}$ 有有界变差，则序列 $\{a_n\}$ 收敛.

(b) 求一个发散的无穷级数 Σa_i，它的元素 a_i 构成一个有有界变差的序列.

(c) 证明以下由戴德肯提出的阿贝尔收敛判别法的推广（参看第 541 页）:

如果 Σa_i 在有限的上下界之间摆动，并且 $\{p_i\}$ 是一个有有界变差且趋于零的序列，则级数 $\Sigma a_i p_i$ 是收敛的.

(d) 证明以下的无穷级数对任何一个固定的实数 x 的收敛性:

(i) $\displaystyle\sum_{n=2}^{\infty} \frac{\sin nx}{\log n}(-1)^n$,

(ii) $\displaystyle\sum_{n=2}^{\infty} \frac{\cos nx}{\log n}(-1)^n$.

14. 讨论以下级数的敛散性:

(a) $\Sigma \dfrac{(-1)^\nu}{\nu}$,

(b) $\Sigma \dfrac{(-1)^\nu \cos(\theta/\nu)}{\nu}$,

(c) $\Sigma \dfrac{\cos \nu\theta}{\nu}$,

(d) $\Sigma \dfrac{\sin \nu\theta}{\nu}$,

(e) $\Sigma \dfrac{(-1)^\nu \cos \nu\theta}{\nu}$,

(f) $\Sigma \dfrac{(-1)^\nu \sin \nu\theta}{\nu}$.

15. 求以下关于 $\log 2$ 的级数

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\frac{1}{6}+-\cdots$$

的重排的和:

(a) $1-\dfrac{1}{2}-\dfrac{1}{4}+\dfrac{1}{3}-\dfrac{1}{6}-\dfrac{1}{8}+\dfrac{1}{5}-\dfrac{1}{10}-\dfrac{1}{12}+--\cdots$,

(b) $1+\dfrac{1}{3}+\dfrac{1}{5}-\dfrac{1}{2}-\dfrac{1}{4}-\dfrac{1}{6}+++---\cdots$.

16. 判断下面的级数是收敛还是发散:

(a) $1+\dfrac{1}{2}-\dfrac{1}{3}+\dfrac{1}{4}+\dfrac{1}{5}-\dfrac{1}{6}+\dfrac{1}{7}+\dfrac{1}{8}-\dfrac{1}{9}++-\cdots$,

(b) $1+\dfrac{1}{2}-\dfrac{2}{3}+\dfrac{1}{4}+\dfrac{1}{5}-\dfrac{2}{6}+\dfrac{1}{7}+\dfrac{1}{8}-\dfrac{2}{9}++-\cdots$.

7.2 节，第 546 页

1. 证明 $\sum_{\nu=2}^{\infty} \frac{1}{\nu(\log \nu)^{\alpha}}$ 当 $\alpha > 1$ 时收敛,而当 $\alpha \leqslant 1$ 时发散.

2. 证明 $\sum_{\nu=3}^{\infty} \frac{1}{\nu \log \nu (\log \log \nu)^{\alpha}}$ 当 $\alpha > 1$ 时收敛,而当 $\alpha \leqslant 1$ 时发散.

3. 如果 n 是任意一个大于 1 的整数,证明

$$\sum_{\nu=1}^{\infty} \frac{a_{\nu}^{n}}{\nu} = \log n,$$

其中 a_{ν}^{n} 定义如下:

$$a_{\nu}^{n} = \begin{cases} 1 & \text{如果 } n \text{ 不是 } \nu \text{ 的因子,} \\ -(n-1) & \text{如果 } n \text{ 是 } \nu \text{ 的因子.} \end{cases}$$

4. 证明 $\sum_{\nu=2}^{\infty} \frac{\log(\nu+1) - \log \nu}{(\log \nu)^{2}}$ 收敛.

5. 证明 $\sum_{\nu=1}^{\infty} \frac{1 \cdot 2 \cdot 3 \cdots \nu}{(\alpha+1)(\alpha+2)\cdots(\alpha+\nu)}$ 当 $\alpha > 1$ 时收敛，而当 $\alpha \leqslant 1$时发散.

*6. 通过与级数$\sum_{\nu=1}^{\infty} \frac{1}{\nu^{\alpha}}$ 比较,证明以下判别法:

如果对于某一个不依赖于 n 的固定数 $\varepsilon > 0$ 和每一个充分大的 n, $\frac{\log(1/|a_n|)}{\log n} > 1 + \varepsilon$, 则级数 Σa_{ν} 绝对收敛;如果对于每一个充分大的 n 和某一个不依赖 n 的数 $\varepsilon > 0$, $\frac{\log(1/|a_n|)}{\log n} < 1 - \varepsilon$, 则级数 Σa_{ν} 不绝对收敛.

7. 证明级数 $\sum_{\nu=1}^{\infty} \left(1 - \frac{1}{\sqrt{\nu}}\right)^{\nu}$ 收敛.

8. 对于什么样的 α 值下面的级数收敛?

(a) $1 - \frac{1}{2^{\alpha}} + \frac{1}{3} - \frac{1}{4^{\alpha}} + \frac{1}{5} - \frac{1}{6^{\alpha}} + - \cdots$.

(b) $1 + \frac{1}{3^{\alpha}} - \frac{1}{2^{\alpha}} + \frac{1}{5^{\alpha}} + \frac{1}{7^{\alpha}} - \frac{1}{4^{\alpha}} + - \cdots$.

9. 通过与级数 $\Sigma \frac{1}{\nu(\log \nu)^{\alpha}}$ 比较,证明以下判别法: 级数 $\Sigma |a_{\nu}|$ 收敛

还是发散取决于对每一个充分大的 n

$$\frac{\log(1/n|a_n|)}{\log\log n}$$

是大于 $1+\varepsilon$ 还是小于 $1-\varepsilon$。

10．从第 6 题的判别法导出根式判别法。

11．证明下面的比较判别法：如果正项级数 Σb_ν 收敛，并且从某一项以后

$$\left|\frac{a_{n+1}}{a_n}\right| < \frac{b_{n+1}}{b_n},$$

则级数 Σa_ν 绝对收敛；如果 Σb_ν 发散，并且从某一项以后

$$\left|\frac{a_{n+1}}{a_n}\right| > \frac{b_{n+1}}{b_n},$$

则级数 Σa_ν 不绝对收敛。

*12．通过与 $\sum_{\nu=1}^{\infty}\frac{1}{\nu^\alpha}$ 比较，证明"汝阿贝 (Raabe)"判别法：

级数 $\Sigma|a_\nu|$ 收敛或发散取决于对每一个充分大的 n 和某一个不依赖于 n 的 $\varepsilon>0$，

$$n\left(\frac{|a_n|}{|a_{n+1}|}-1\right)$$

是大于 $1+\varepsilon$ 还是小于 $1-\varepsilon$。

13．通过与 $\Sigma\frac{1}{\nu(\log\nu)^\alpha}$ 比较，证明以下判别法：级数 $\Sigma|a_\nu|$ 收敛或发散取决于对每一个充分大的 n

$$n\log n\left(\frac{|a_n|}{|a_{n+1}|}-1-\frac{1}{n}\right)$$

是大于 $1+\varepsilon$ 还是小于 $1-\varepsilon$。

14．证明高斯 (Gauss) 判别法：

如果

$$\frac{|a_n|}{|a_{n+1}|}=1+\frac{\mu}{n}+\frac{R_n}{n^{1+\varepsilon}},$$

其中 $|R_n|$ 是有界的，并且 $\varepsilon>0$ 不依赖于 n，则 $\Sigma|a_\nu|$ 当 $\mu>1$ 时收敛，而当 $\mu\leqslant1$ 时发散。

15．判别下面的"超几何"级数的敛散性：

(a) $\frac{\alpha}{\beta}+\frac{\alpha(\alpha+1)}{\beta(\beta+1)}+\frac{\alpha(\alpha+1)(\alpha+2)}{\beta(\beta+1)(\beta+2)}+\cdots$.

(b) $1+\frac{\alpha\cdot\beta}{1\cdot\gamma}+\frac{\alpha(\alpha+1)\cdot\beta(\beta+1)}{1\cdot2\cdot\gamma(\gamma+1)}$

$$+\frac{\alpha(\alpha+1)(\alpha+2)\cdot\beta(\beta+1)(\beta+2)}{1\cdot2\cdot3\cdot\gamma(\gamma+1)(\gamma+2)}+\cdots.$$

7.4 节，第 555 页

1. 序列 $f_n(x)$, $n=1,2,\cdots$ 在区间 $0\leqslant x\leqslant1$ 上用等式

$$f_0(x)\equiv1,\qquad f_n(x)=\sqrt{xf_{n-1}(x)}$$

定义．

(a) 证明在区间 $0\leqslant x\leqslant1$ 上序列收敛到一个连续的极限．

*(b) 证明这收敛性是一致的．

*2. 设 $f_0(x)$ 在区间 $0\leqslant x\leqslant a$ 上连续，函数序列 $f_n(x)$ 用

$$f_n(x)=\int_0^x f_{n-1}(t)dt,\qquad n=1,2,\cdots$$

定义．证明这序列在区间 $0\leqslant x\leqslant a$ 上一致收敛到 0．

*3. 设 $f_n(x)$, $n=1,2,\cdots$ 是在区间 $a\leqslant x\leqslant b$ 上具有连续导数的函数序列．证明，如果 $f_n(x)$ 在区间的每一点处收敛，同时不等式 $|f'_n(x)|<M$（其中 M 是一个常数）为一切 n 和 x 值所满足，则收敛是一致的．

4. (a) 证明级数 $\sum_{\nu=1}^{\infty}\frac{1}{\nu^x}$ 对任一固定的数 $\varepsilon>0$，对 $x\geqslant1+\varepsilon$ 一致收敛．

(b) 证明求导得出的级数 $-\Sigma\frac{\log\nu}{\nu^x}$ 对固定的正数 ε，对 $x\geqslant1+\varepsilon$ 一致收敛．

*5. 证明级数 $\Sigma\frac{\cos\nu x}{\nu^\alpha}$，$\alpha>0$，对任意小的正值 ε，对 $\varepsilon\leqslant x\leqslant2\pi-\varepsilon$ 一致收敛．

6. 级数

$$\frac{x-1}{x+1}+\frac{1}{3}\left(\frac{x-1}{x+1}\right)^3+\frac{1}{5}\left(\frac{x-1}{x+1}\right)^5+\cdots$$

当 ε,N 为固定的正数时，对 $\varepsilon\leqslant x\leqslant N$ 一致收敛．

7. 求下列级数收敛的范围：

(a) $\Sigma x^{\nu!}$，　　　(b) $\Sigma\frac{(\nu!)^2x^\nu}{(2\nu)!}$，

(c) $\Sigma \frac{a^\nu}{\nu^x}$ $(a < 1)$, (d) $\Sigma \frac{a^\nu}{\nu^x}$ $(a > 1)$,

(e) $\Sigma \frac{\log \nu}{\nu^x}$, (f) $\Sigma \frac{x^\nu}{1 - x^\nu}$,

*8. 证明如果狄里克莱级数 $\Sigma \frac{a_\nu}{\nu^x}$ 对 $x = x_0$ 收敛，则它对任意的 $x > x_0$ 收敛；如果它对 $x = x_0$ 发散，则它对任意的 $x < x_0$ 发散. 因此存在一个"收敛的横坐标"使得对任一个大于它的 x 值级数收敛，而对任一个小于它的 x 值级数发散.

9. 如果 $\Sigma \frac{a_\nu}{\nu^x}$ 对 $x = x_0$ 收敛，则求导得出的级数 $-\Sigma \frac{a_\nu \log \nu}{\nu^x}$ 对任意的 $x > x_0$ 收敛。

7.5 节，第 566 页

1. 如果幂级数 $\Sigma a_n x^n$ 的收敛区间是 $|x| < \rho$，同时 $\Sigma b_n x^n$ 的收敛区间是 $|x| < \rho'$，这里 $\rho < \rho'$，那末 $\Sigma (a_n + b_n) x^n$ 的收敛区间是什么?

2. 如果 $a_\nu > 0$ 并且 Σa_ν 收敛，则

$$\lim_{x \to 1-0} \Sigma a_\nu x^\nu = \Sigma a_\nu.$$

3. 如果 $a_\nu > 0$ 并且 Σa_ν 发散，则

$$\lim_{x \to 1-0} \Sigma a_\nu x^\nu = \infty.$$

*4. 证明阿贝尔定理：

如果 $\Sigma a_\nu X^\nu$ 收敛，则 $\Sigma a_\nu x^\nu$ 对 $0 \leqslant x \leqslant X$ 一致收敛。

*5. 如果 $\Sigma a_\nu X^\nu$ 收敛，则 $\lim\limits_{x \to X-0} \Sigma a_\nu x^\nu = \Sigma a_\nu X^\nu$.

*6 根据幂级数的乘法证明

(a) $e^x e^y = e^{x+y}$. (b) $\sin 2x = 2 \sin x \cos x$.

7. 利用二项式级数，计算 $\sqrt{2}$ 到四位小数.

8. 设 a_n 为任一个实数序列，S 是 a_n 的所有的极限点的集合. 我们用 $p = \overline{\lim}\, a_n$ 表示 S 的最小上界 p. 证明幂级数 $\sum\limits_{n=0}^{\infty} c_n x^n$ 当 $|x| < \rho$ 时收敛，当 $|x| > \rho$ 时发散，其中

$$\rho = \frac{1}{\overline{\lim} \sqrt[n]{|c_n|}}.$$

附录，第 582 页

1．证明关于 $\sqrt{(1-x)}$ 的幂级数当 $x=1$ 时仍然收敛.

2．证明对于每一个正数 ε，在区间 $0\leqslant x\leqslant 1$ 上存在一个表示 $\sqrt{(1-x)}$ 的 x 的多项式，它与 $\sqrt{(1-x)}$ 的误差小于 ε.

3．在问题 2 里设 $x=1-t^2$，证明对每一个正数 ε，存在一个 t 的多项式，它在区间 $-1\leqslant t\leqslant 1$ 上表示 $|t|$，它与 $|t|$ 的误差小于 ε.

4．(a) 证明如果 $f(x)$ 在 $a\leqslant x\leqslant b$ 上连续，则对每一个 $\varepsilon>0$，存在一个多边形函数 $\varphi(x)$（即，一个连续函数，它的图形由有限个相交成角的直线段组成）使得对区间里的每一个 x 有 $|f(x)-\varphi(x)|<\varepsilon$.

(b) 证明每一个多边形函数 $\varphi(x)$ 能够用一个和函数 $\varphi(x)=a+bx+\Sigma c_i|x-x_i|$ 表示，其中 x_i 是角点的横坐标.

5．**外尔斯特拉斯逼近定理.** 在上面论述的基础上证明，如果 $f(x)$ 在 $a\leqslant x\leqslant b$ 上连续，则对每一个正数 ε 存在一个多项式 $P(x)$ 使得对区间 $a\leqslant x\leqslant b$ 上所有的 x 值有 $|f(x)-P(x)|<\varepsilon$.

提示：用形如 $(x-x_r)+|x-x_r|$ 的线性组合去逼近 $f(x)$.

6．证明以下的无穷乘积收敛：

(a) $\displaystyle\prod_{n=1}^{\infty}\left(1+\left(\frac{1}{2}\right)^{2n}\right)$,

(b) $\displaystyle\prod_{n=2}^{\infty}\frac{n^3-1}{n^3+1}$,

(c) $\displaystyle\prod_{n=1}^{\infty}\left(1-\frac{z^2}{n!}\right)\qquad(|z|<1)$.

7．用课文中的方法证明 $\displaystyle\prod_{n=1}^{\infty}\left(1+\frac{1}{n}\right)$ 发散.

8．对 $|x|<1$ 证明恒等式

$$\prod_{\nu=1}^{\infty}(1+x^{2^\nu})=\frac{1}{1-x}.$$

*9．在十进制所表示的自然数中，考虑所有数字中间没有 9 的数．证明这些数的倒数和收敛.

10．(a) 证明，对于 $s>1$,

$$1-\frac{1}{2^s}+\frac{1}{3^s}-\frac{1}{4^s}+\cdots=(1-\alpha^{1-s})\zeta(s),$$

其中 $\zeta(s)$ 是在第 587 页中所定义的 Zeta 函数。

(b) 用这个恒等式去证明 $\lim\limits_{s\to 1+0}(s-1)\zeta(s)=1$。

11．关于收敛的积分判别法　设 $f(x)$ 是正的并且当 $x\geqslant 1$ 时是递减的.

(a) 证明无穷积分

$\int_1^{\infty} f(x)dx$ 和无穷级数 $\sum\limits_{k=1}^{\infty} f(k)$ 同时收敛或同时发散.

(b) 证明不论在哪种情况下，极限

$$\lim_{n\to\infty}\left(\int_1^n f(x)dx-\sum_{k=1}^{n} f(k)\right)$$

都存在.

(c) 应用这个判别法证明级数

$$\sum_{n=2}^{\infty}\frac{1}{n\log^{\alpha} n}$$

当 $\alpha>1$ 时收敛，当 $\alpha\leqslant 1$ 时发散。

第八章 三角级数

用幂级数表示的函数，或者如拉格朗日所称呼的“解析函数”，在分析中的确起着一种中心的作用．但是，由于这类解析函数在许多实例中太受限制，因此，下述事实对整个数学以及对大量应用来说，是一个颇为重大的事件．这就是，傅立叶在他的《热的分析理论》[1]一书中，注意到并且用许多例阐明了，具有常系数 a_ν，b_ν 的形如

$$f(x)=\frac{a_0}{2}+\sum_{\nu=1}^{\infty}(a_\nu\cos\nu x+b_\nu\sin\nu x) \tag{1}$$

的收敛的三角级数能够表示一类广泛的“任意”函数 $f(x)$．这类函数基本上包括了每一个特别重要的函数，不管是用机械作图的方法从几何上定义的，还是用其它方法定义的，甚至具有跳跃间断的函数，或者在不同区间上服从不同的形成规律的函数，都能用(1)式表示出来．

在傅立叶的引人注目的发现之后不久，“傅立叶级数”就被认为不仅是物理学和力学的最强有力的工具，而且同样也恰是许多漂亮的纯数学结果的一个丰富的源泉．在 1820 到 1830 年间，哥西，特别是狄锐西勒对于傅立叶的有点启发的和不完善的论证提供了一个坚实的基础，使得这个课题如像它的重要性一样易于接受．

不管由三角级数所表示的函数的“任意性”如何，因为级数的每一项都有 2π 这个周期，所以它们本来就受着 2π 这一周期性条件的限制．但是，我们不久就会看到，这个限制是非本质的，只要我们仅在一个有限的区间内考虑函数，我们就能容易地把它开拓

1) 见英译本：Joseph Fourier 著“The Analytical Theory of Heat”, Dover 1955 年再版发行．

为周期函数.

本章对傅立叶级数的理论提供一个初步介绍，而不作更进一步的仔细研究.

在对周期函数作了某些初步的讨论之后，我们将证明主要定理,以便对于一类广泛的函数建立其三角级数展开的合理性.

在以后的几节中,我们将讨论一点儿较高深的附加的课题,例如傅立叶级数的一致和绝对收敛性以及任意连续函数的多项式近似. 在附录中,我们将讨论伯努利的多项式理论及其应用.

8.1. 周 期 函 数

a. 一般说明. 函数的周期开拓

由于函数 $\sin nx$ 和 $\cos nx$ 是以 2π 为公共周期的 x 的周期函数;所以形如(1)的任何有限和或收敛的无限和也是以 2π 为周期的周期函数. 我们现在作一些关于周期函数的一般性的观察，以扩充第四章第 360 页的内容.

周期为 T 的函数 $f(x)$ 的周期性,由对所有 x 的值[1]都成立的等式

$$f(x+T)=f(x) \tag{2a}$$

来表示. 具有周期 T 意味着 $f(x)$ 也具有周期 $\pm T$，$\pm 2T$，$\cdots$，$\pm mT$，$\cdots$，并且对所有整数 m 都有

$$f(x\pm mT)=f(x). \tag{2b}$$

在特殊情况下，$f(x)$ 也可以偶而有一个较短的周期. 例如，函数 $\sin(4\pi x/T)$ 不仅有周期 T,而且同样有较小的周期 $T/2$.

如同我们已在第四章第 361 页看到过的，定义在一个闭区间 $a\leqslant x\leqslant b$ 内的函数 $f(x)$，能够开拓成对 x 的所有值具有周期

---

1) 在表示周期函数时，把自变量 x 看作是一个圆周上的点以代替在一条直线上的点,这常常是方便的. 对于一个周期为 2π 的函数 $f(x)$，我们考虑一个单位圆的圆心角 x，它介于一个任意起始半径与圆周上的一个动点的半径之间; 于是 $f(x)$ 的周期性就意味着圆周上的每一个点刚好对应着函数的一个值,尽管角度 x 本身还可有 2π 的整数倍的差别.

$T = b - a$ 的周期函数，它在原来的区间 $a \leqslant x \leqslant b$ 以外的依次相邻的长度为 T 的区间内是按下列周期性关系

$$f(x + nT) = f(x), \qquad n = \pm 1, \pm 2, \cdots \tag{2c}$$

定义的.

在长度为 T 的区间的端点 $x = a + nT = b + (n-1)T$ 上，由开拓所得的函数既不是唯一确定的，也未必连续. 我们必须允许 $f(x)$ 在点 $x = \xi$ 具有跳跃间断性；函数 $f(x)$ 不论在 ξ 的哪一边都是连续的，但在点 ξ 本身却未必是确定的或连续的.

因此，下述符号和 $f(\xi)$ 的定义遍及本章. 我们用

$$f(\xi + 0) = \lim_{\varepsilon \to 0} f(\xi + \varepsilon^2), \tag{3a}$$

$$f(\xi - 0) = \lim_{\varepsilon \to 0} f(\xi - \varepsilon^2) \tag{3b}$$

来表示 $f(x)$ 在 $x = \xi$ 处的右极限和左极限；在 f 的间断点 ξ 上，不管 $f(\xi)$ 原来取怎样的值，都**定义**它取**平均值**

$$f(\xi) = \frac{1}{2}[f(\xi + 0) + f(\xi - 0)], \tag{4}$$

这样作是方便的.

借助这个约定，就可将我们原来的函数从闭区间 $a \leqslant x \leqslant b$ 上周期地开拓到 x 的全部值，即使在 $f(a) \neq f(b)$ 的情况中也可以. 我们需要注意的只是在跳跃间断的点上 $f(x)$ 的值，特别是，如果这种跳跃间断是由原来定义的 $f(a)$ 和 $f(b)$ 的值不相同而引起的；为了定义周期的开拓，我们必须采用平均值 $\frac{1}{2}[f(a)+f(b)]$ 来代替 $f(a)$ 和 $f(b)$ 的值.

b. 一个周期上的积分

周期函数 $f(x)$ 的图形，在相当于一个周期的任何两个相邻的区间内显然具有相同的形状. 这隐含着下述重要的事实：对于周期为 T 的周期函数 $f(x)$ 和任意 a 都有

$$\int_{-a}^{T-a} f(x)dx = \int_0^T f(x)dx, \tag{5}$$

或者说：在长度为T的一个周期的区间上，周期函数的积分总是具有相同的值，而不管区间位于何处.

要证明这个事实，我们只需注意，根据等式 $f(\xi-T)=f(\xi)$，替换 $x=\xi-T$ 对于任何 α，β 都给出

$$\int_{\alpha}^{\beta} f(x)dx=\int_{\alpha+T}^{\beta+T} f(\xi)d\xi=\int_{\alpha+T}^{\beta+T} f(x)dx.$$

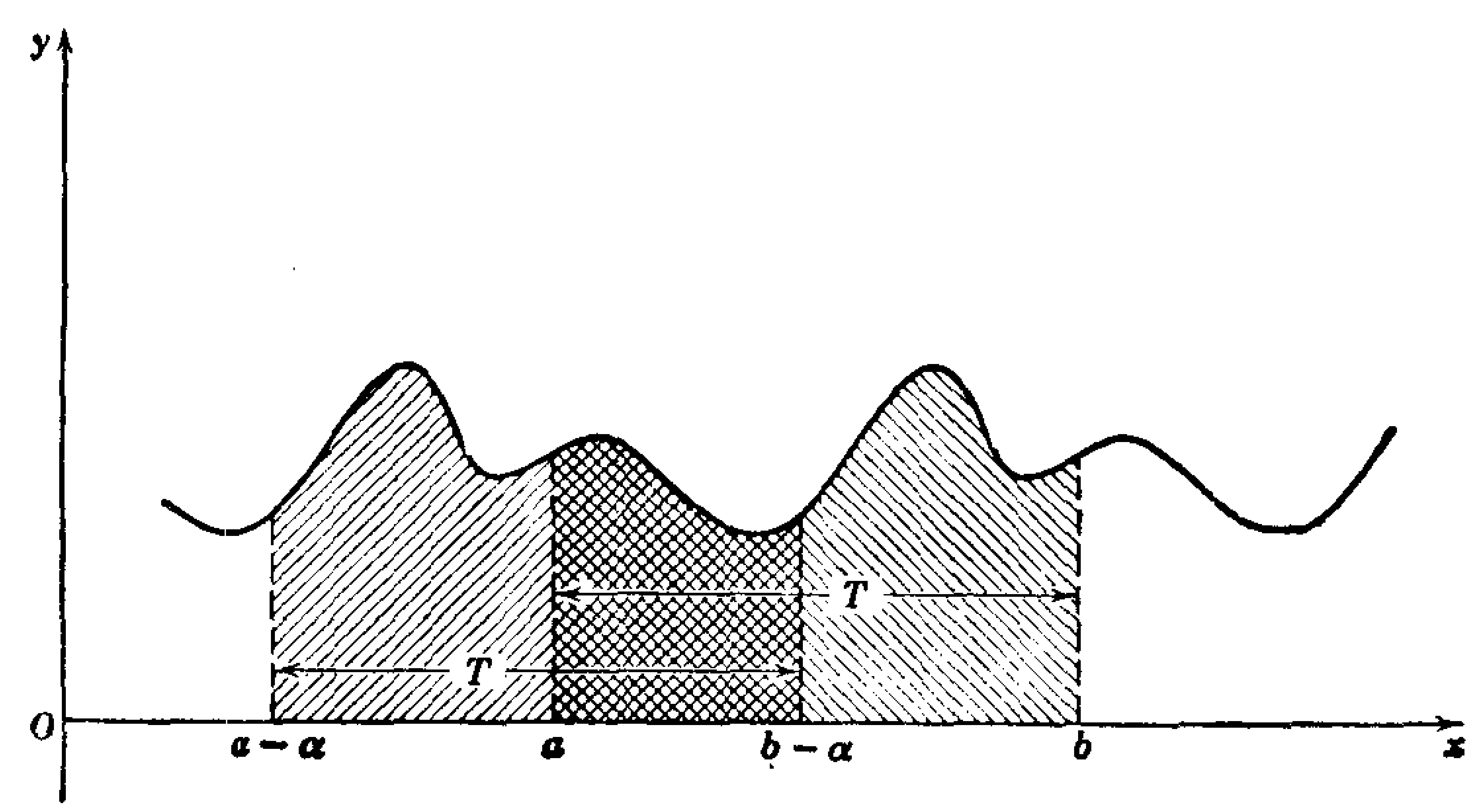


图 8.1 一个整周期上的积分的图解

特别是，当 $\alpha=-a$ 和 $\beta=0$ 时，有

$$\int_{-a}^{0} f(x)dx=\int_{T-a}^{T} f(x)dx,$$

因此

$$\begin{aligned}\int_{-a}^{T-a} f(x)dx&=\int_{-a}^{0} f(x)dx+\int_{0}^{T-a} f(x)dx\\&=\int_{T-a}^{T} f(x)dx+\int_{0}^{T-a} f(x)dx\\&=\int_{0}^{T} f(x)dx,\end{aligned}$$

正如所述．联想起积分的几何意义，就由图 8.1 可看出结论显然成立.

c. 谐振

我们将从最简单的周期函数出发去构造出最一般的函数；这

些最简单的周期函数是 $a\sin\omega x$ 和 $a\cos\omega x$，或更一般地是 $a\sin\omega(x-\xi)$ 和 $a\cos\omega(x-\xi)$，其中 $a(\geqslant 0)$，$\omega(>0)$ 和 ξ 都是常数. 这些函数代表"正弦振动"或"简谐振动"(或振荡)[1]. 振动的周期为 $T=\frac{2\pi}{\omega}$. 数 ω 称为振动的圆频率或角频率[2]；因为 $\frac{1}{T}=\frac{\omega}{2\pi}$ 是单位时间内振动的次数或频率，ω 就是 2π 时间内振动的次数. 数 a 称为振动的振幅；它代表函数 $a\sin\omega(x-\xi)$ 或 $a\cos(x-\xi)$ 的最大值，因为正弦和余弦两者都有最大值 1. 数 $\omega(x-\xi)$ 称为位相，同时数 $\omega\xi$ 称为相位移或相移.

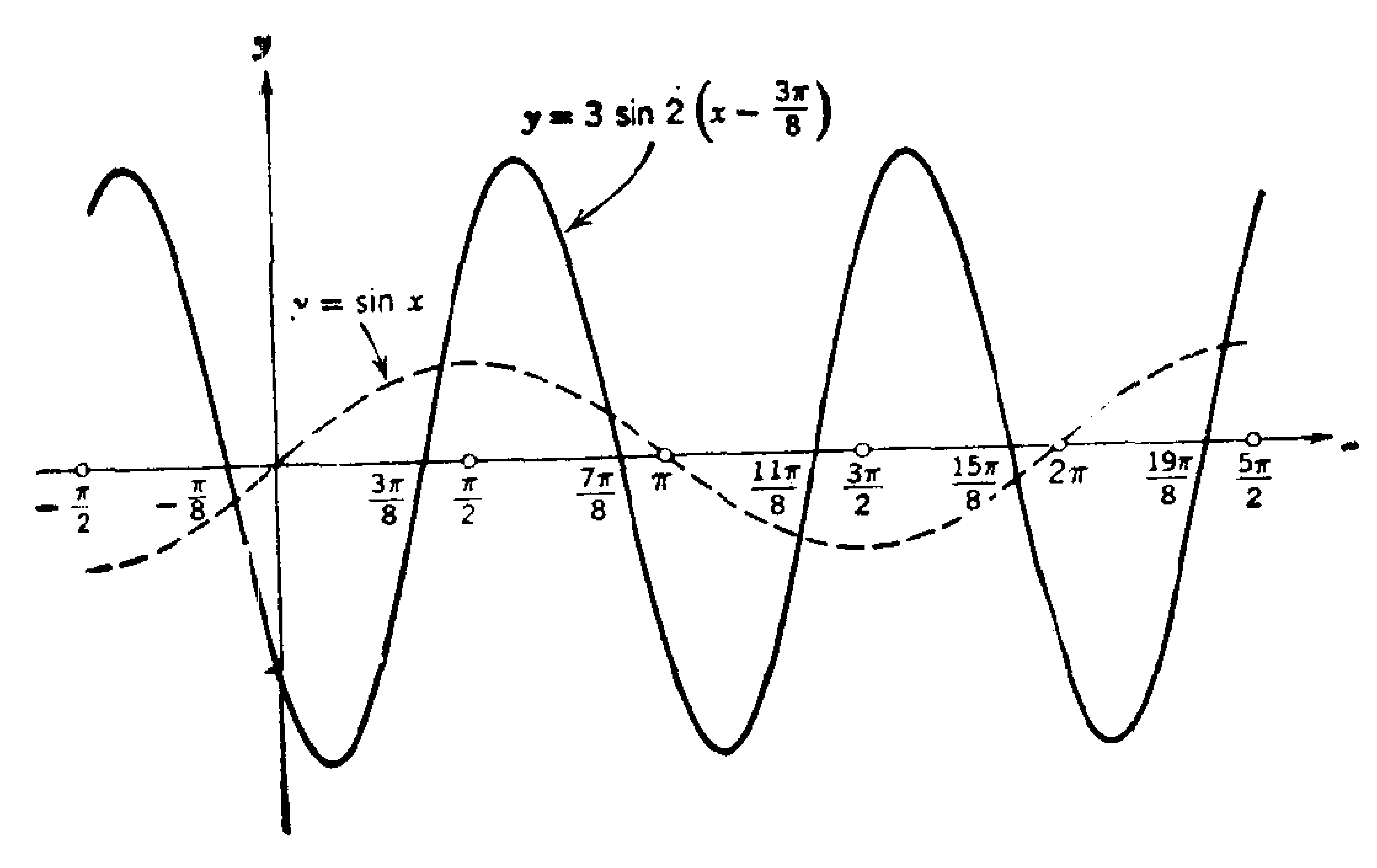


图 8.2　正弦振动

沿 x 轴以 $1:\omega$ 的比例同时沿 y 轴以 $a:1$ 的比例伸缩变换正弦曲线，再沿着 x 轴的正方向将曲线平移一个距离 ξ（参见图 8.2），我们就得到函数 $a\sin\omega(x-\xi)$ 的图形.

根据三角函数的加法公式，我们也能分别用 $\alpha\cos\omega x+\beta\sin\omega x$ 以及 $\beta\cos\omega x-\alpha\sin\omega x$ 表示谐振，其中 $\alpha=-a\sin\omega\xi$

1) 单独看这些表达式的任何一个(对所有的 a 和 ξ)，都表示所有正弦振动的集合；这两个表达式是等价的，因为 $a\sin\omega(x-\xi)=a\cos\omega\left[x-\left(\xi+\frac{\pi}{2\omega}\right)\right]$.

2) 注意我们对频率和圆频率之间的区别.

和 $\beta = a\cos\omega\xi$. 反过来，每个形如 $\alpha\cos\omega x + \beta\sin\omega x$ 的函数表示一个由方程 $\alpha = -a\sin\omega\xi$, $\beta = a\cos\omega\xi$ 给出的，具有振幅 $a = \sqrt{\alpha^2 + \beta^2}$ 和相位移 $\omega\xi$ 的正弦振动 $a\sin\omega(x - \xi)$. 利用表达式 $\alpha\cos\omega x + \beta\sin\omega x$，我们立刻能够将两个或两个以上具有相同圆频率 ω 的这种函数的和，写成具有圆频率 ω 的另外的一个振动.

正如早就看到的，周期函数是在我们希望用参数表示闭曲线的时候出现的. 自然，它们能够用来表示由圆运动产生的现象，譬如说，用来表示一个和飞轮一样的周期地重复的过程；此外，它们也与所有振动现象相联系.

8.2 谐振的迭加

a. 谐波. 三角多项式

虽然许多振动是纯正弦的（参看第 428 页），但是大部分周期运动具有较复杂的特征，它们是由若干正弦振动"迭加"而成的. 在数学上，直线上的点的运动，它的坐标 x 为时间 t 的函数，可由上述类型的若干纯周期函数的和函数所给定. 这时，函数的谐波分量就被迭加起来了（也就是它们的纵坐标被加起来了）. 在这一迭加中，我们假定迭加振动的圆频率（当然周期也是）均不相同，因为如上所述，具有相同圆频率的两个正弦振动的迭加所产生的是具有相同的圆频率的另一个正弦振动.

对于具有不同圆频率 ω_1 和 ω_2 的两个正弦振动的迭加，存在两个基本上不同的可能性，这取决于 ω_1/ω_2 是否有理数，或者，如我们说过的，取决于这两个频率是可公度的，还是不可公度的.

作为第一种情况的例，我们假定第二个圆频率是第一个圆频率的两倍：$\omega_2 = 2\omega_1$. 于是，第二个振动的周期为第一个振动的周期的一半，即 $2\pi/2\omega_1 = T_2 = T_1/2$，因此，它不仅具有周期 T_2，而且也具有两倍于 T_2 的周期 T_1，因为，函数本身经周期 T_1 之后又周而复始；由此可知，由迭加所形成的函数必须同样具有周期 T_1. 圆频率为第一个振动的两倍，周期为其一半的第二个振动称为第一

个振动(基波)的一次谐波 (first harmonic).

如果我们引入具有圆频率 $\omega_3 = 3\omega_1$ 的另一振动,相应的叙述也是正确的. 于是函数 $\sin 3\omega_1 x$ 本身经 $2\pi/\omega_1 = T_1$ 之后,必又周而复始. 这一振动称为给定振动的二次谐波. 类似地,我们能够考虑具有圆频率 $\omega_4 = 4\omega_1$, $\omega_5 = 5\omega_1$, $\cdots$, $\omega_n = n\omega_1$ 的三次,四次,$\cdots$,$(n-1)$ 次谐波,并可具有我们所希望的任何相位移. 每一个这种谐波本身都必然按周期 $\frac{2\pi}{\omega_1} = T_1$ 周而复始. 因而,由一批振动迭加得到的每个函数都是具有周期 $2\pi/\omega = T_1$ 的周期函数,这些振动的每一个是给定基频 ω_1 的一个谐波. 把具有从基波到 $(n-1)$ 次谐波的圆频率的振动都迭加起来,我们就得到一个形式为三角多项式的周期函数

$$s_n(x) = \frac{a_0}{2} + \sum_{\nu=1}^{n} (a_\nu \cos \nu\omega x + b_\nu \sin \nu\omega x). \qquad (6)$$

(常数 $\frac{a_0}{2}$ 不影响周期性,加上它是为了以后的方便.) 由于这个函数包含 $2n+1$ 个任意常数 a_ν, b_ν, 所以我们能作成各种曲线,与原来的正弦曲线大不一样. 图 8.3 到 8.5 是图示说明.

"谐波"这个术语是指声学[1]说的,在声学里一个频率为 ω 的基频振动相当于某个一定高低的纯音,第一,第二,第三等等谐波对应于一系列基音的谐音,也就是八度音加上第五音程,二倍八度音等等.

在一般情况下,对于其圆频率具有有理比的那些振动的迭加而言,圆频率均可表示成公共基频的整数倍.

然而对于具有不可公度的圆频率 ω_1 和 ω_2 的两个振动的迭加,呈现出不同的现象. 这里正弦振动的迭加不再是周期的了. 我们不进行详细的讨论,仅指出这种函数具有一种"近似周期"的特性,或者如我们所说的,它们是殆周期的 (almost periodic).

1) 在声学里也采用泛音这一术语.

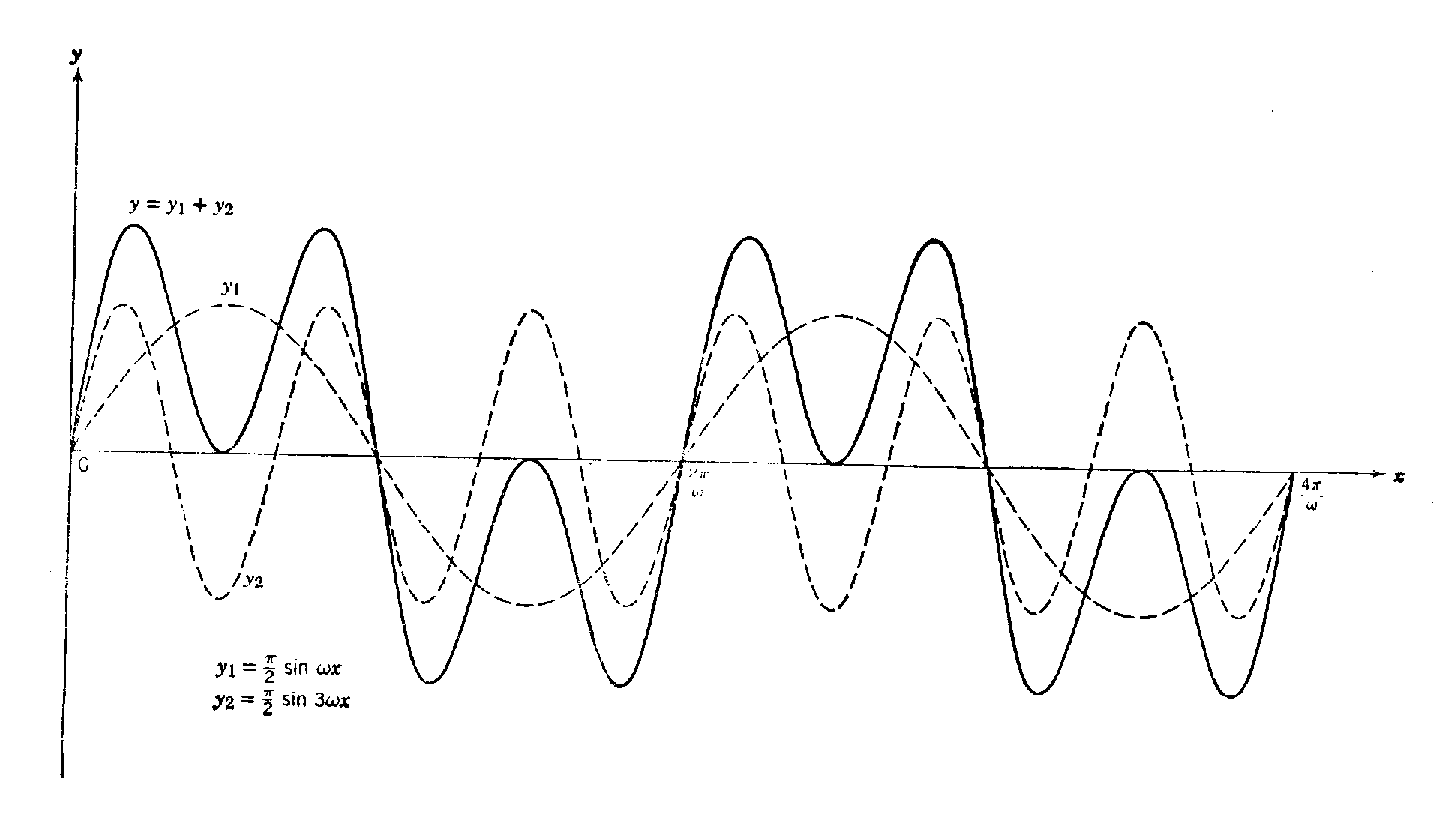


图 8.3　振动的组合(图的比例对应于采取 $\omega = 1.$)

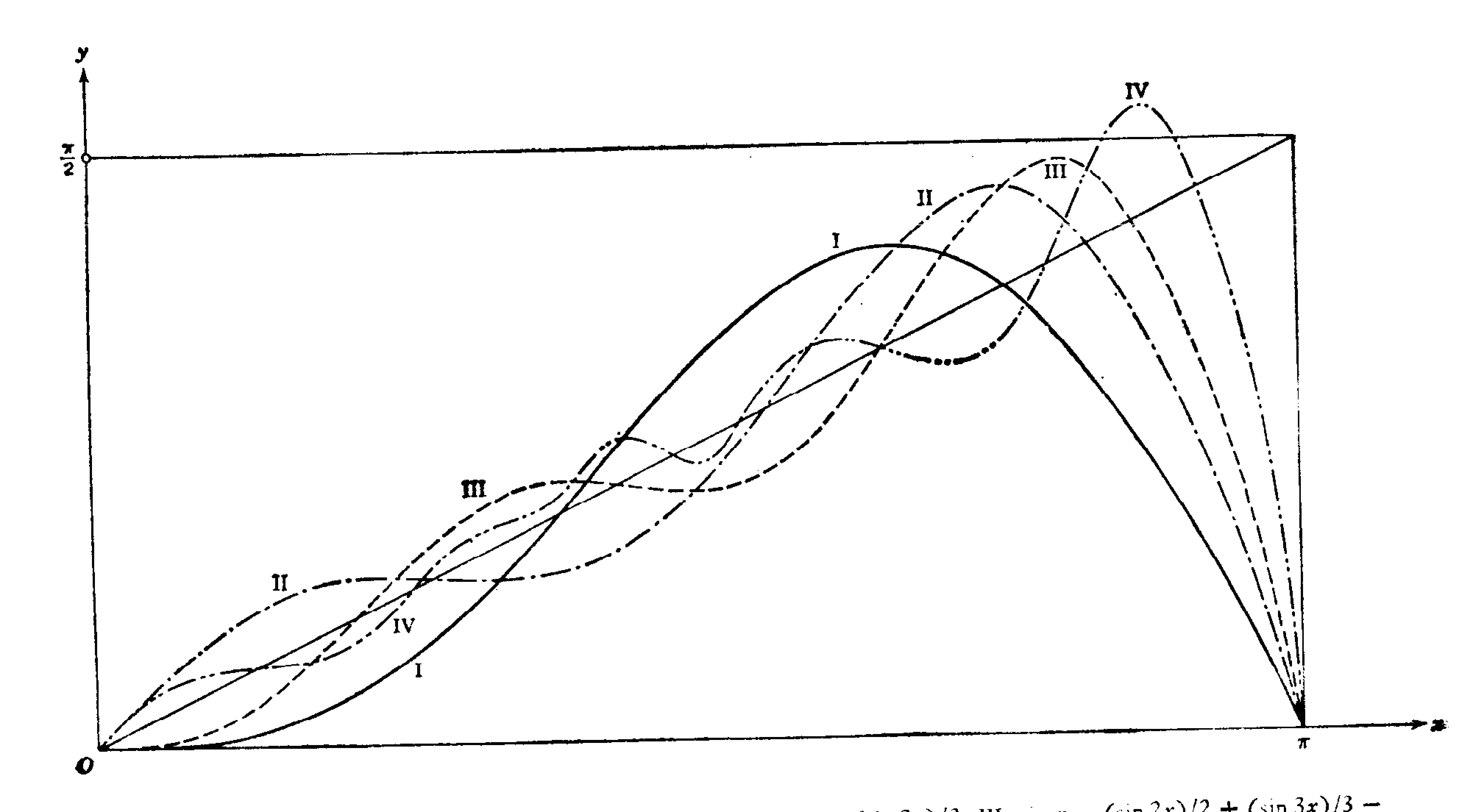


图 8.4　振动的组合. I: $\sin x-(\sin 2x)/2$; II: $\sin x-(\sin 2x)/2+(\sin 3x)/3$; III: $\sin x-(\sin 2x)/2+(\sin 3x)/3-(\sin 4x)/4$; IV: $\sin x-(\sin 2x)/2+(\sin 3x)/3-(\sin 4x)/4+(\sin 5x)/5-(\sin 6x)/6+(\sin 7x)/7$

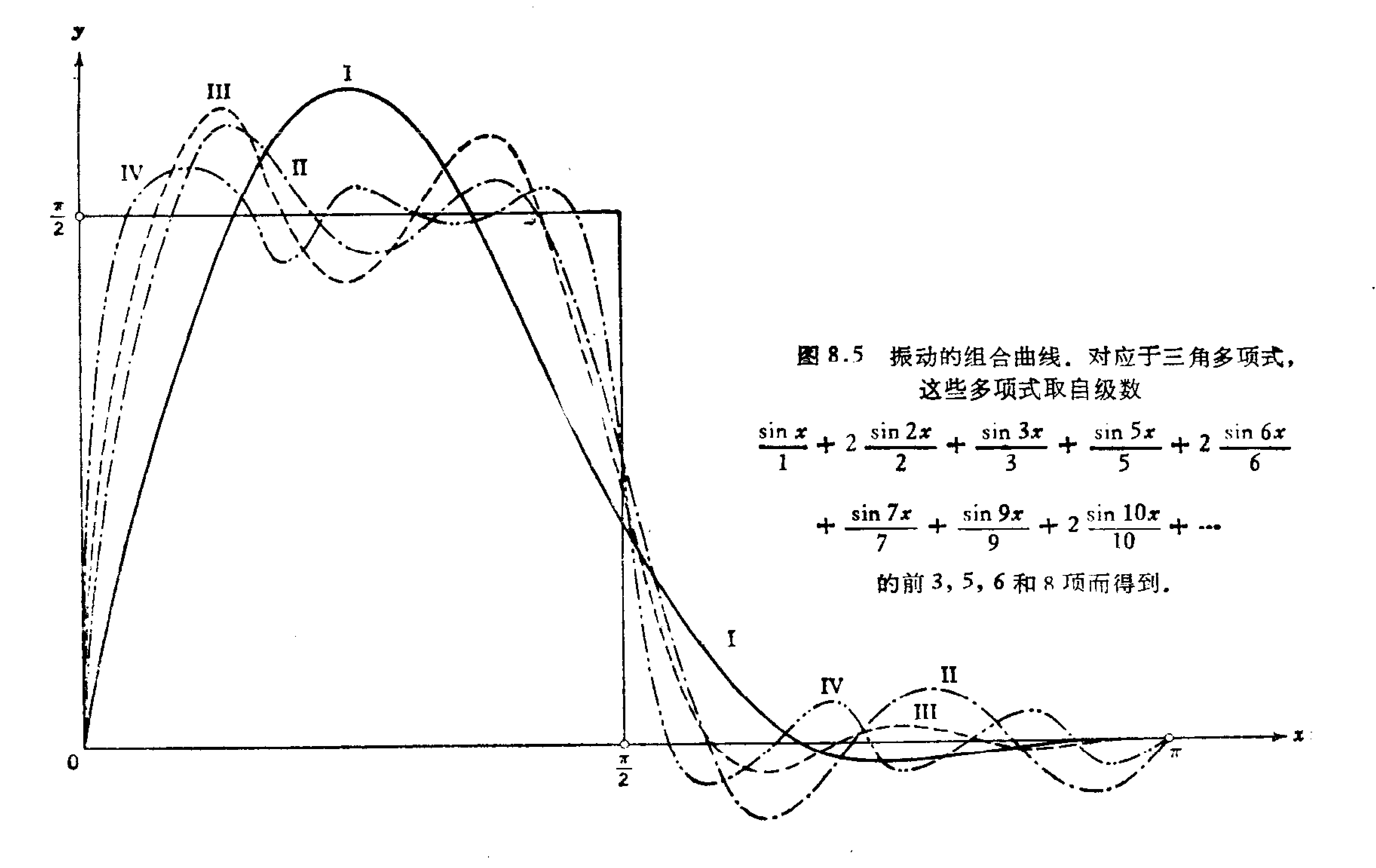


图 8.5　振动的组合曲线．对应于三角多项式，这些多项式取自级数

$$\frac{\sin x}{1}+2\frac{\sin 2x}{2}+\frac{\sin 3x}{3}+\frac{\sin 5x}{5}+2\frac{\sin 6x}{6}+\frac{\sin 7x}{7}+\frac{\sin 9x}{9}+2\frac{\sin 10x}{10}+\cdots$$

的前 3，5，6 和 8 项而得到．

*b. 拍

对正弦振动迭加的最后一个注记牵涉到所谓拍(beat)的现象。如果我们把两个振幅各为1，圆频率分别为 ω_1 和 ω_2 的振动迭加起来，并且为了简单起见，假定两者取相同的 ξ 值(见第604页，推广到任意相位的工作留给读者)，那末我们就要考虑函数

$$y = \sin\omega_1 x + \sin\omega_2 x \qquad (\omega_1 > \omega_2 > 0).$$

根据一个众所周知的三角公式我们有

$$y = 2\cos\left[\frac{1}{2}(\omega_1 - \omega_2)x\right]\sin\left[\frac{1}{2}(\omega_1 + \omega_2)x\right].$$

这个方程表示一个现象，我们把它描述如下：

我们有一个圆频率为 $\frac{1}{2}(\omega_1 + \omega_2)$ 和周期为 $4\pi/(\omega_1 + \omega_2)$ 的振动，这个振动不具有一个常幅，而具有由表达式

$$2\cos\left[\frac{1}{2}(\omega_1 - \omega_2)x\right]$$

给出的一个变"幅"，这个变幅以一个较长的周期 $4\pi/(\omega_1 - \omega_2)$ 变化。当两个圆频率 ω_1 和 ω_2 相当大而其差 $(\omega_1 - \omega_2)$ 又比较小时，

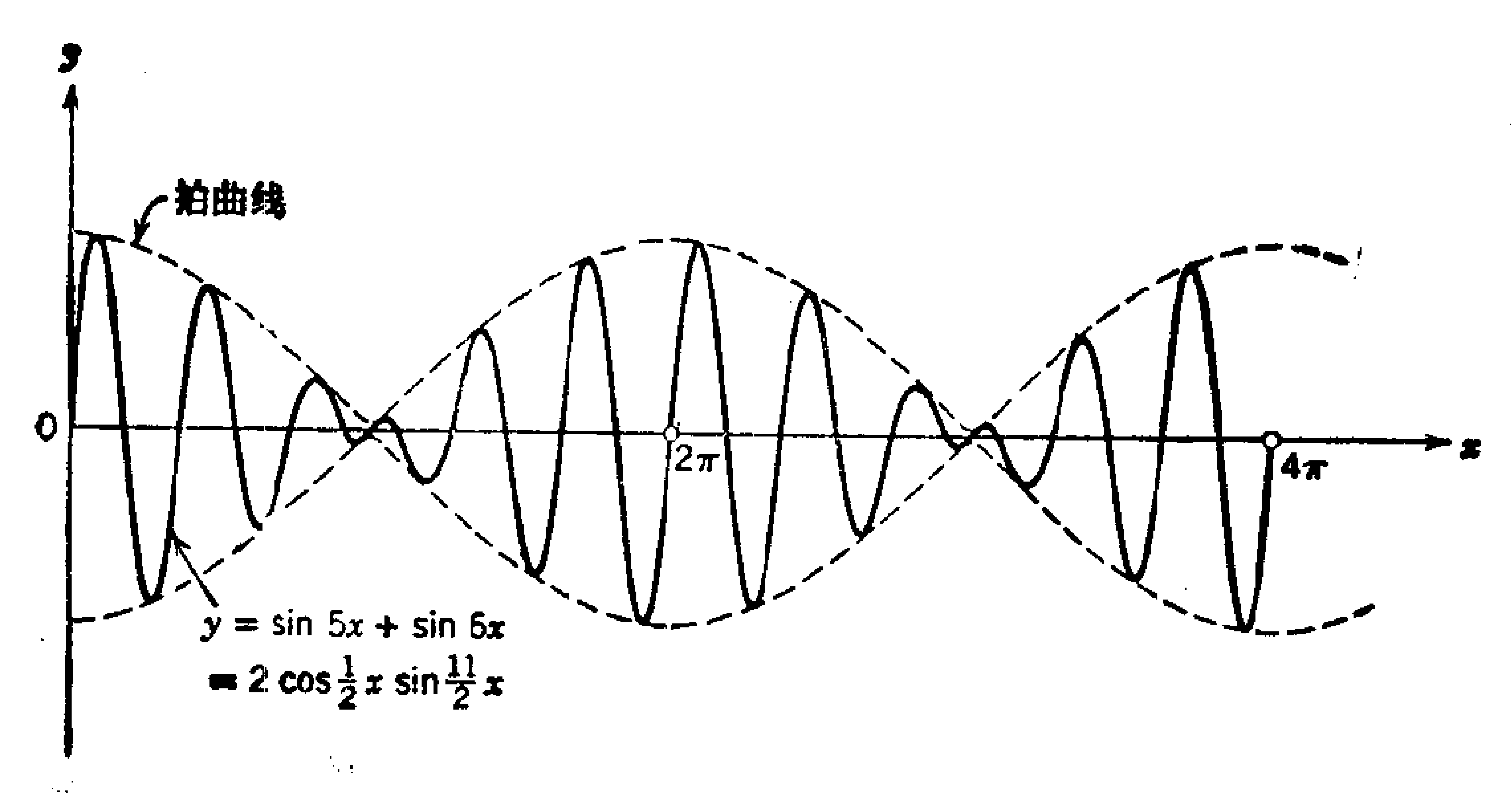


图 8.6 拍

这个描述特别有用。这时，振幅 $2\cos\left[\frac{1}{2}(\omega_1 - \omega_2)x\right]$ 作周期性

缓慢变化，其周期为 $\frac{4\pi}{\omega_1-\omega_2}$，比振动的周期 $\frac{4\pi}{\omega_1+\omega_2}$ 要长。这些振幅的有节奏的变化称为**拍**．每个人都是熟悉声学和电子学中的这种现象的．通常，在无线电传播中，圆频率 ω_1 和 ω_2 远远超过耳朵所能听到的范围，然而其差 $\omega_1-\omega_2$ 却在耳朵可听到的音调范围之内．因此，这个拍引起一个可听到的音，而原来的振动依然存在于耳朵的觉察之外．

在图 8.6 中所描述的，就是拍的一个例．

8.3 复数表示法

a. 一般说明

三角函数的运算通常根据欧拉的关系式

$$\cos\theta+i\sin\theta=e^{i\theta}$$

或

$$\cos\theta=\frac{1}{2}(e^{i\theta}+e^{-i\theta}), \tag{7a}$$

$$\sin\theta=\frac{1}{2i}(e^{i\theta}-e^{-i\theta}), \tag{7b}$$

利用复数来简化(对照第七章第 578 页)．因此，我们能够通过复数量 $e^{i\omega x}$，$e^{-i\omega x}$，或 $ae^{i\omega(x-\xi)}$，$ae^{-i\omega(x-\xi)}$ 来表示正弦振动，其中 a，ω 和 $\omega\xi$ 分别是振幅，圆频率和相位移．当然最终实际的振动不难从复数表达式中经分离实部和虚部而得到．

复数表示法的方便之一在于，关于时间 x 求导数时可以把 i 视为实常数，而对指数函数求导；由正弦函数和余弦函数的求导公式得出

$$\begin{aligned}&\frac{d}{dx}a[\cos\omega(x-\xi)+i\sin\omega(x-\xi)]\\&\quad=a\omega[-\sin\omega(x-\xi)+i\cos\omega(x-\xi)]\\&\quad=ia\omega[\cos\omega(x-\xi)+i\sin\omega(x-\xi)],\end{aligned}$$

这可以简明地写成形式

$$\frac{d}{dx} ae^{i\omega(x-\xi)} = ia\omega e^{i\omega(x-\xi)}. \qquad (8)$$

一个复值函数 $\gamma(x)$，譬如说 $\gamma(x) = p(x) + iq(x)$，其积分自然地由下式定义

$$\int \gamma(x)dx = \int p(x)dx + i\int q(x)dx.$$

因此，对 $n \neq 0$ 有

$$\int e^{inx}dx = \int \cos nx dx + i\int \sin nx dx$$

$$= \frac{1}{n}\sin nx - \frac{i}{n}\cos nx = \frac{1}{in}e^{inx}.$$

特别地，对任意整数 n 我们有

$$\int_{-\pi}^{\pi} e^{inx}dx = \begin{cases} 0, & \text{当 } n \neq 0, \\ 2\pi, & \text{当 } n = 0. \end{cases}$$

更一般地，如果我们记住 $e^{inx}e^{-imx} = e^{i(n-m)x}$，就可见对于任何整数 m,n，我们有

$$\int_{-\pi}^{\pi} e^{inx}e^{-imx}dx = \begin{cases} 0, & \text{当 } n \neq m, \\ 2\pi, & \text{当 } n = m. \end{cases} \qquad (9)$$

这些关系式只不过是三角函数之间的正交关系的简明表达式（见第 292 页）.

*b. 交流电上的应用

我们用一个重要的例来说明这些思想，在例中用 t 代替 x 来表示自变量，时间.

我们考虑一个具有电阻 R 和电感 L 的电路，在其上加一个电动势（电压）E. 在直流电中，电压是常数，电流 I 由欧姆定律 $E = RI$ 给出. 然而对于交流电而言，E 是时间 t 的函数，从而 I 也是时间 t 的函数，并且欧姆定律取推广了的形式（见第 673 页）

$$E - L\frac{dI}{dt} = RI. \qquad (10)$$

我们考虑一个外电动势 E，它是以 ω 为圆频率的正弦波，由 $e\cos\omega t$ 或 $e\sin\omega t$

给出，我们把二者形式地组合成复的形式

$$E = \varepsilon e^{i\omega t} = \varepsilon\cos\omega t + i\varepsilon\sin\omega t,$$

其中 ε 表示振幅．对振幅也允许用复的值

$$\varepsilon = |\varepsilon| e^{-i\eta},$$

这常常是有用的；于是

$$E = |\varepsilon| e^{i(\omega t-\eta)} = |\varepsilon|\{\cos(\omega t-\eta) + i\sin(\omega t-\eta)\}.$$

我们可以把 i 看作好象是一个实参数一样，而对这个"复电压" E 和相应的复电流 I 进行运算．于是复的量 E 和 I 之间的复的关系的意义在于，相应于电动势 $\varepsilon\cos\omega t$ 的电流是 I 的实部，而相应于电动势 $\varepsilon\sin\omega t$ 的电流是 I 的虚部．复电流由形如

$$I = \alpha e^{i\omega t} = \alpha(\cos\omega t + i\sin\omega t)$$

的表达式给出，它也是具有圆频率 ω 的正弦波．于是 I 的导数形式由

$$\frac{dI}{dt} = i\alpha\omega e^{i\omega t}$$

$$= \alpha\omega(-\sin\omega t + i\cos\omega t) = i\omega I$$

给出．将这些量代入欧姆定律的推广了的形式（10）并且除以因子 $e^{i\omega t}$，我们得到方程

$$\varepsilon - \alpha L i\omega = R\alpha$$

或

$$\alpha = \frac{\varepsilon}{R + i\omega L},$$

以及

$$E = (R + i\omega L)I = WI.$$

如果我们称量

$$W = R + i\omega L$$

为电路的复电阻，我们就可以把上面的方程看作在复数形式下交流电的欧姆定律．于是欧姆定律如像直流电一样：电流等于电压被电阻除．

令 $w = |W|$，把复电阻 W 写成

$$W = we^{i\delta} = w\cos\delta + iw\sin\delta,$$

式中

$$|W| = w = \sqrt{(R^2 + L^2\omega^2)}, \qquad \tan\delta = \frac{\omega L}{R},$$

我们得到

$$I = \frac{\varepsilon}{w} e^{i(\omega t - \delta)} = \frac{E}{W}.$$

按照这个公式，电流和电压具有相同的周期(及圆频率)；对于实的 ε，电流的振幅 α 同电动势的振幅 ε 按照

$$\alpha = \frac{\varepsilon}{w}$$

互相联系着，此外，电流与电压之间存在一个相位差．电流与电压不在同一个时间达到最大值，而是落后了一段时间 δ/ω．当然，对于最小值也同样如此．在电工学中，量 $w = \sqrt{R^2 + L^2\omega^2}$ 常常称圆频率为 ω 时电路的**阻抗**或**交流电阻**；相位移通常用度表示，有时称为**滞后**．

如果振幅 ε 是形如

$$\varepsilon = |\varepsilon| e^{-i\eta}$$

的复数，除 η 是一个附加的相移之外，这时欧姆定律的形式没有本质上的变化，我们有

$$E = |\varepsilon| e^{i(\omega t - \eta)},$$

$$I = \frac{E}{W} = \frac{|\varepsilon|}{|W|} e^{i\omega t} e^{-i(\delta + \eta)}.$$

c. 三角多项式的复数表示法

形如

$$s_n(x) = \frac{1}{2} a_0 + \sum_{\nu=1}^{n} (a_\nu \cos \nu x + b_\nu \sin \nu x) \tag{11}$$

(为简便起见，我们取 $\omega = 1$) 的复合振动用代换

$$\cos \nu x = \frac{1}{2}(e^{i\nu x} + e^{-i\nu x}), \quad \sin \nu x = -\frac{1}{2} i (e^{i\nu x} - e^{-i\nu x}),$$

可化为复数形式．这时表达式取得较简单的形式

$$s_n(x) = \sum_{\nu=-n}^{n} \alpha_\nu e^{i\nu x}, \tag{12}$$

式中复数 α_ν 按照等式

$$
\begin{cases}
\alpha_\nu = \frac{1}{2}(a_\nu - ib_\nu), \\
\alpha_{-\nu} = \frac{1}{2}(a_\nu + ib_\nu), \qquad \text{对于 } \nu = 1, 2, \cdots, n, \\
\alpha_0 = \frac{1}{2} a_0,
\end{cases}
\tag{13a}
$$

与实数 a_0, a_ν 和 b_ν 相联系. 对于 a_ν 和 b_ν 解这些关系式,我们求得

$$
\begin{cases}
a_\nu = \alpha_\nu + \alpha_{-\nu}, \\
b_\nu = i(\alpha_\nu - \alpha_{-\nu})
\end{cases}
\tag{13b}
$$

(包括 $\nu = 0$ 的情况).

反过来,我们可把形如

$$\sum_{\nu=-n}^{n} \alpha_\nu e^{i\nu x}$$

的任何表达式看作一个表示振动迭加的函数，这些振动都写成了复的形式. 当且仅当 $\alpha_\nu + \alpha_{-\nu}$ 是实的且 $\alpha_\nu - \alpha_{-\nu}$ 是纯虚的时;也就是当 α_ν 和 $\alpha_{-\nu}$ 是共轭复数时,这一迭加的结果才是实的

d. 一个三角公式

作为复数表示法的一个应用,我们证明下面这个恒等式:

$$
\sigma_n(\alpha) = \frac{1}{2} + \cos\alpha + \cos 2\alpha + \cdots + \cos n\alpha
$$

$$
= \frac{\sin\left(n + \frac{1}{2}\right)\alpha}{2\sin\frac{1}{2}\alpha}, \tag{14}
$$

这在本章后续部分是需要的. 这个公式只有在 $\sin\frac{1}{2}\alpha \neq 0$ 时,即当 $\alpha \neq 0, \pm 2\pi, \pm 4\pi, \cdots$ 时才有意义. 但是,这个公式一旦对 $\sin\frac{1}{2}\alpha \neq 0$ 建立起来之后,我们就可断言,表达式$\left[\sin\left(n+\frac{1}{2}\right)\alpha\right/$

$2\sin\frac{1}{2}\alpha\Big]$ 对一切 α 都是 α 的连续函数，只要我们将公式在例外点的值定义为 $\sigma_n(\alpha)$ 在这些点的值，即定义为 $n+\frac{1}{2}$，就可以了.

为了证明这一公式，我们用余弦的指数表达式来代替余弦函数[见公式(13a)，当 $a_\nu=1$，$b_\nu=0$]：

$$\sigma_n(\alpha)=\frac{1}{2}\sum_{\nu=-n}^{n}e^{i\nu\alpha}.$$

右边是一个公比为 $q=e^{i\alpha}=\cos\alpha+i\sin\alpha$ 的几何级数. 因此仅当 $\cos\alpha=1$，$\sin\alpha=0$ 时，也就是说，如果 α 取例外的值之一：0，$\pm 2\pi$，$\pm 4\pi,\cdots$ 时，q 才能取值 1. 对于所有其它的值 α，由普通求和公式得到

$$\sigma_n(\alpha)=\frac{1}{2}e^{-in\alpha}\frac{1-q^{2n+1}}{1-q}$$

$$=\frac{1}{2}\frac{e^{-in\alpha}-e^{(n+1)i\alpha}}{1-e^{i\alpha}}.$$

用 $e^{-i\alpha/2}$ 乘分子和分母，我们得到上述的公式：

$$\sigma_n(\alpha)=\frac{\sin\left(n+\frac{1}{2}\right)\alpha}{2\sin\frac{1}{2}\alpha}.$$

在 $0\leqslant t\leqslant\pi$ 上积分 $\sigma_n(t)$，我们得到一个有用结果，它不依赖于 n：

$$\int_0^\pi\frac{\sin\left(n+\frac{1}{2}\right)t}{2\sin\frac{1}{2}t}dt=\int_0^\pi\left(\frac{1}{2}+\sum_{\nu=1}^{n}\cos\nu t\right)dt=\frac{1}{2}\pi\qquad(15)$$

这是因为级数的每一项的积分都为零的缘故.

8.4 傅立叶级数

a. 傅立叶系数

阶数为 n 的三角多项式

$$f(x)=s_n(x)=\frac{1}{2}a_0+\sum_{\nu=1}^{n}(a_\nu\cos\nu x+b_\nu\sin\nu x) \qquad (16)$$

依赖于 $(2n+1)$ 个系数 a_ν 和 b_ν. 值得注意的是，这些"傅立叶系数"能够通过级数的和函数 $f(x)$ 的值由下面的公式简单地表示出来

$$a_\mu=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos\mu x dx, \quad b_\mu=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin\mu x dx. \qquad (17)$$

证明是容易的，只需用 $\cos\mu x$ 或 $\sin\mu x$ 乘(16)而后积分就行了. 由正交关系(见第 292 页)可直接推出这些表达式，因为只有 $\nu=\mu$ 的项才不为零.

写成复数形式

$$f(x)=s_n(x)=\sum_{\nu=-n}^{n}\alpha_\nu e^{i\nu x}, \qquad (16a)$$

$$a_\nu=\alpha_\nu+\alpha_{-\nu}; \qquad b_\nu=i(\alpha_\nu-\alpha_{-\nu}),$$

对于复的傅立叶系数，正如在第 612 页上根据复的正交关系(9)已经看到的，相应的表达式是

$$\alpha_\nu=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(x)e^{-i\nu x}dx. \qquad (17a)$$

附带指出，在(16)式的常数项 $\frac{1}{2}a_0$ 的记法中，因子 $\frac{1}{2}$ 只是为了使公式(17)在 $\nu=0$ 时成立.

现在我们引出傅立叶级数的主要定理，方法是提出这样一个自然的问题，即令傅立叶多项式(16)的阶数 n 趋于无穷时，它是否能够表示函数 $f(x)$，这个函数除了以 2π 为周期而外本质上是任意的.

在下文中我们的主要结果就是：任意的一个分段连续并有分段连续的一阶和二阶导数的周期函数 $f(x)$，能够用一个无穷的"傅立叶级数"

$$f(x)=\frac{a_0}{2}+\sum_{\nu=1}^{\infty}(a_\nu\cos\nu x+b_\nu\sin\nu x)$$

来表示，或写成复数形式

$$f(x)=\sum_{\nu=-\infty}^{\infty}\alpha_\nu e^{i\nu x}$$

其中系数由(17)和(17a)给出.

b. 基本引理

我们首先回忆在一个区间内逐段或分段连续的函数是这样定义的，就是一个函数在区间内除去有限多个跳跃间断点外是连续的.

我们再进一步回忆，一个周期函数 $f(x)$ 在间断点上的值定义为从两边求出的极限值的平均，如同早已约定的那样 [第 602 页等式(4)].

如果我们能把整个区间分为有限个子区间，使得 f, f', f'' 在每个开子区间内是连续的并在端点趋于确定的极限，则函数 $f(x)$ 是分段连续的，并具有分段连续的一阶和二阶导数.

证明主要定理的关键是下面的简单的事实:

引理. 如果一个函数 $k(x)$ 及其一阶导数 $k'(x)$ 在区间 $a\leqslant x\leqslant b$ 上是分段连续的，则当 $\lambda\to\infty$ 时，积分

$$k_\lambda=\int_a^b k(x)\sin\lambda x dx$$

趋于零.

证明. 我们用分部积分法来证明这个引理. 假定 k 和 k' 在 $a\leqslant x\leqslant b$ 上是连续的，我们有

$$\begin{aligned}k_\lambda&=\int_a^b k(x)\sin\lambda x dx\\&=\frac{1}{\lambda}\left[k(a)\cos\lambda a-k(b)\cos\lambda b+\int_a^b k'(x)\cos\lambda x dx\right];\quad(18)\end{aligned}$$

当 λ 增加时，右边显然趋于零. 如果 $k(x)$ 或 $k'(x)$ 在区间里有跳

跃间断点 ξ，那末我们就用这些点 ξ 把区间剖分成部分区间，把我们的论证应用到这些部分区间上，然后把所得结果加起来，仍然得到所要证的结果.

我们指出(省略证明)，无需任何有关导数 $k'(x)$ 存在性的假定，而只需 k 的分段连续性，引理事实上仍是正确的. 在这些微弱的条件下，证明要依赖于当 $\lambda \neq 0$ 时，函数 $\sin \lambda x$ 在长度为 π/λ 的相邻区间里交替地为正和为负. 对于大的 λ 值，由于函数 $k(x)$ 的连续性，使得相邻区间上的积分值，几乎彼此抵消了.

c. $\int_0^\infty \frac{\sin z}{z} dz = \frac{\pi}{2}$ 的证明

作为引理的一个应用，我们来求积分

$$I = \int_0^\infty \frac{\sin z}{z} dz \tag{19}$$

的值. 这个广义积分是用关系式

$$I = \lim_{M \to \infty} I_M$$

来定义的，式中

$$I_M = \int_0^M \frac{\sin z}{z} dz.$$

广义积分 I 的收敛性，也就是当 $M \to \infty$ 时，I_M 的极限的存在性，已经在第 332 页中证明过了. 收敛性的证明是以分部积分为根据的，这里可以重述一下. 譬如说，设 $0 < M < N$，我们有

$$\begin{aligned} |I_N - I_M| &= \left| \int_M^N \frac{\sin z}{z} dz \right| \\ &= \left| \frac{-\cos z}{z} \Big|_M^N + \int_M^N \frac{\cos z}{z^2} dz \right| \\ &\leqslant \frac{1}{M} + \frac{1}{N} + \int_M^N \frac{dz}{z^2} = \frac{2}{M}. \end{aligned} \tag{20}$$

因为当 M 和 N 都充分大时，I_N 和 I_M 的差任意小，于是由柯西的收敛判别法就保证了 $I = \lim_{M \to \infty} I_M$ 的存在性. 此外，在(20)中令 N 趋向无穷，我们就找到 I_M 逼近它的极限 I 的速率的一个估计:

$$|I - I_M| \leqslant \frac{2}{M}. \tag{20a}$$

我们能够将 I 的表达式用另一种方式写成为一个固定的有限区间上的积分的极限．令 p 是一个任意的正数；对于 $M = \lambda p$ 代换 $z = \lambda x$, $dz = \lambda dx$ 表明

$$I_{\lambda p} = \int_0^{\lambda p} \frac{\sin z}{z} dz = \int_0^p \frac{\sin \lambda x}{x} dx.$$

由于对固定的正 p 当 $\lambda \to \infty$ 时，$\lambda p \to \infty$，我们显然有

$$I = \lim_{\lambda \to \infty} \int_0^p \frac{\sin \lambda x}{x} dx,$$

并且从(20a)更明显地有

$$\left| I - \int_0^p \frac{\sin \lambda x}{x} dx \right| < \frac{2}{\lambda p}.$$

因此，对任何正的 p，表达式

$$\int_0^p \frac{\sin \lambda x}{x} dx$$

当 $\lambda \to \infty$ 时逼近同一个值 I；此外，对 p 是一致收敛的，只要我们限制 p 的值大于某一个固定的正数 P．于是，当 $\lambda > 2/P\varepsilon$ 时，积分和极限 I 之间的差确实是小于 ε 的．

我们现在运用第 618 页的引理于函数

$$k(x) = \frac{1}{x} - \frac{1}{2\sin(x/2)}.$$

如果我们定义 $k(0) = 0$，函数 $k(x)$ 就对于 $0 \leqslant x < 2\pi$ 是连续的并有连续的一阶导数(见第 490 页)．因此，我们的引理说明了，只要 $0 \leqslant p < 2\pi$，当 $\lambda \to \infty$ 时，

$$\int_0^p \sin \lambda x \left(\frac{1}{x} - \frac{1}{2\sin(x/2)} \right) dx$$

就趋向零．此外，根据(18)，对于 $0 \leqslant p \leqslant \pi$，收敛是一致的，因为 $|k(x)|$ 和 $|k'(x)|$ 在区间 $0 \leqslant x \leqslant \pi$ 上是有界的．对区间 $0 < p < 2\pi$ 上的任何 p，从我们前面的结果得到

$$\lim_{\lambda\to\infty}\int_0^p \frac{\sin\lambda x}{2\sin(x/2)}\,dx = I,$$

而且对于 $P\leqslant p\leqslant\pi$ 中的 p，收敛性也是一致的，其中 P 是一个固定的正数.

现在对 $p=\pi$ 和 $\lambda=n+\frac{1}{2}$（其中 n 是一个整数）我们已经计算了这个积分的值[见第 616 页公式(15)]，并求得了它有同 n 无关的值 $\pi/2$. 令 λ 通过形如 $\lambda=n+\frac{1}{2}$ 的值趋向无穷，我们就求得 I 的值 $\frac{\pi}{2}$:

$$\int_0^\infty \frac{\sin z}{z}\,dz = \frac{\pi}{2}. \tag{21}$$

此外，我们已经证明了

$$\lim_{\lambda\to\infty}\int_0^p \frac{\sin\lambda x}{x}\,dx = \frac{\pi}{2}, \tag{21a}$$

当 $P\leqslant p$ 时收敛是一致的，这里 P 是一个固定的正数，并且

$$\lim_{\lambda\to\infty}\int_0^p \frac{\sin\lambda x}{2\sin(x/2)}\,dx = \frac{\pi}{2}, \tag{21b}$$

这里对于 $P\leqslant p\leqslant\pi$ 收敛是一致的.

d. 函数 $\phi(x)=x$ 的傅立叶展式

我们上述的结果直接引出两个相关的分段线性周期函数 $\varphi(x)$，$\chi(x)$ 的傅立叶展式，它们在区间 $-\pi<x<\pi$ 上，用下式来定义（见图 8.7 和 8.8）:

$$\varphi(x)=x$$

及

$$\chi(x)=\begin{cases}\pi-x, & \text{当 } x>0,\\ 0, & \text{当 } x=0,\\ -\pi-x, & \text{当 } x<0.\end{cases} \tag{22}$$

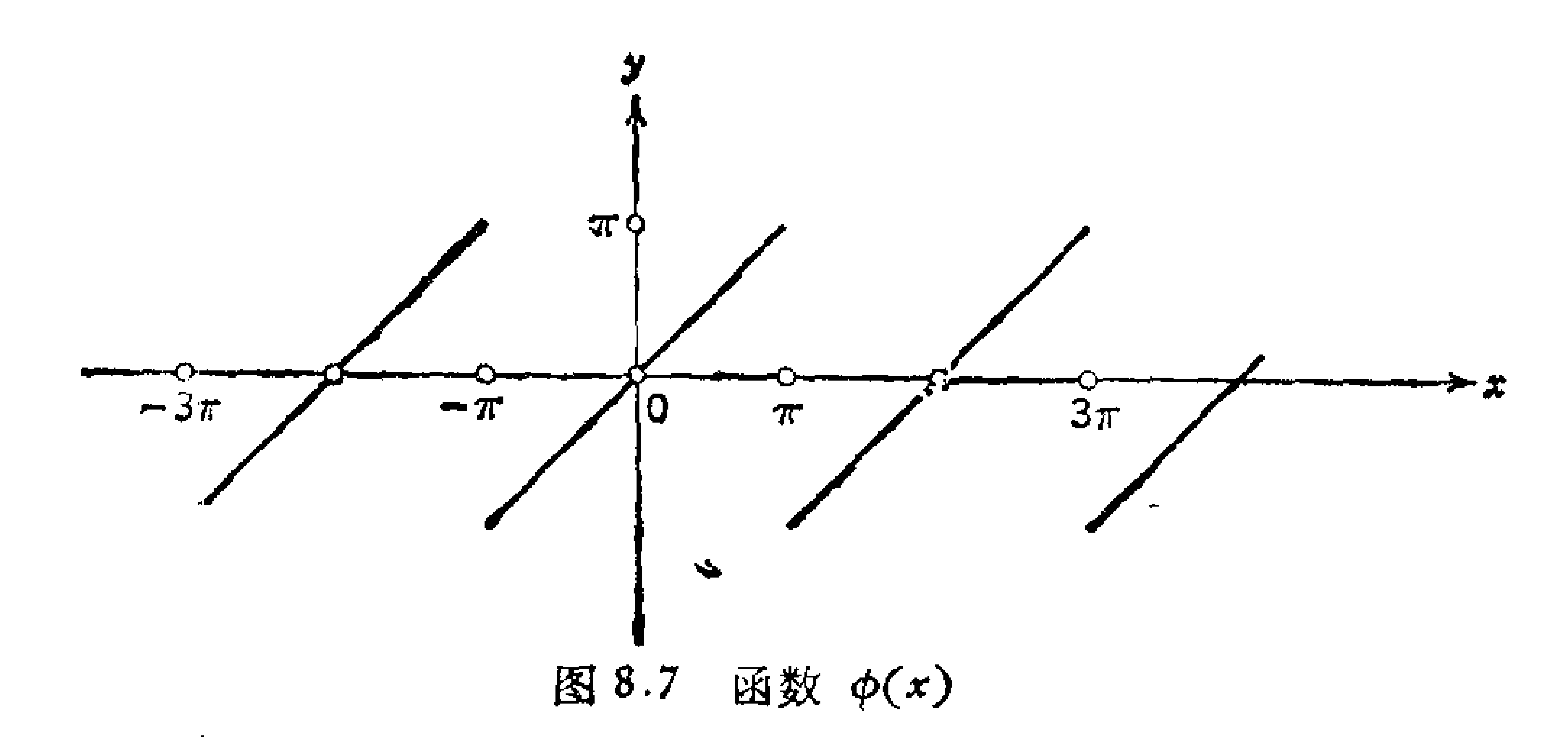


图 8.7 函数 $\phi(x)$

把这些函数周期性地开拓到区间 $-\pi < x < \pi$ 之外，第一个函数 ϕ 在端点具有跳跃间断性，而 $\chi(x)$ 在 $x = 0$ 有一个 2π 的跳跃. 显然，周期性地开拓出来的这两个函数按照

$$\chi(x) = \phi(\pi - x)$$

互相关联着.

对于 $\lambda = n + \frac{1}{2}$ 和 $p = x$，通过极限过程 $n \to \infty$，从第 615 页上的公式(14)和第 621 页上的公式(21b)，直接得到 $\chi(x)$ 在 $0 < x \leqslant \pi$ 上的傅立叶展式. 于是我们求出傅立叶级数

$$\chi(x) = 2\left(\sin x + \frac{\sin 2x}{2} + \frac{\sin 3x}{3} + \cdots\right). \tag{23a}$$

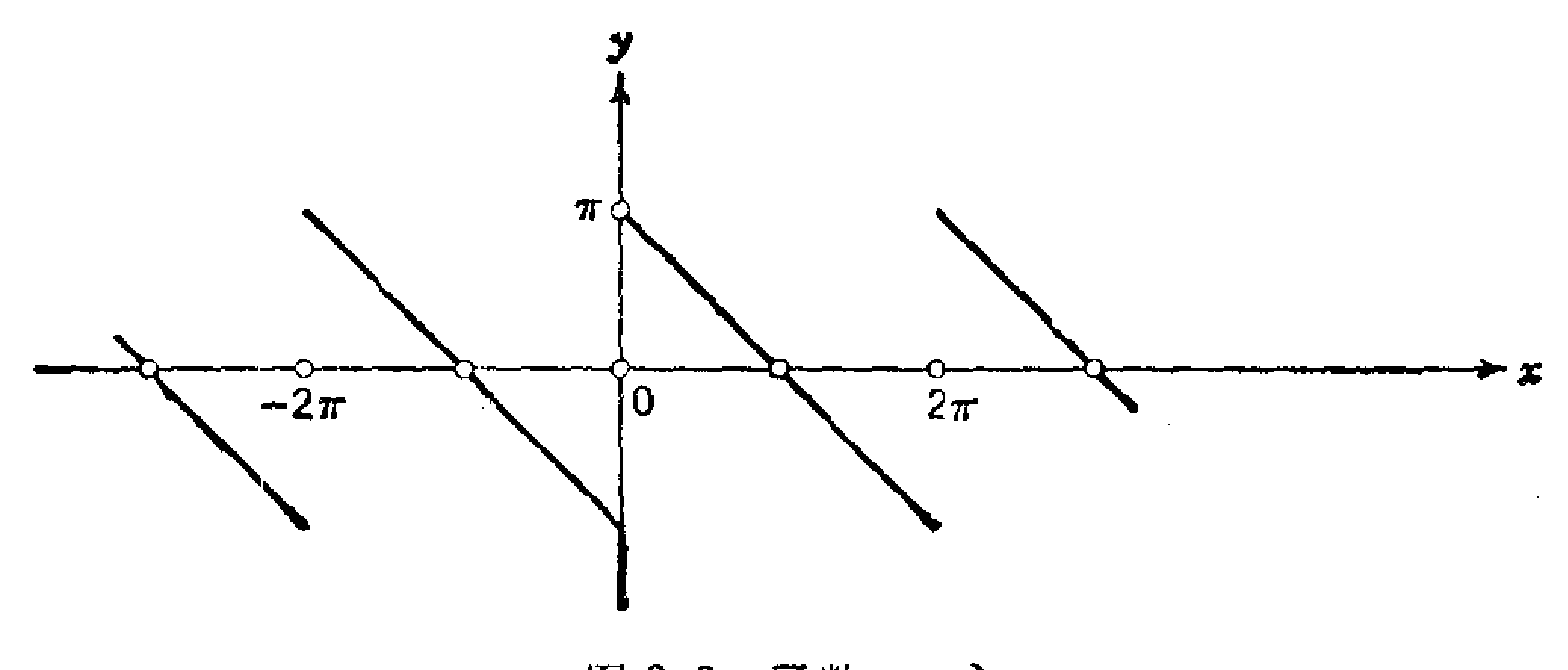


图 8.8 函数 $\chi(x)$

对于 $-\pi \leqslant x < 0$ 也同样适用，因为两边同是 x 的奇函数. 对于任意小的正数 ε，级数在 $\varepsilon < |x| \leqslant \pi$ 上是一致收敛的. 在 $x = 0$ 时级数的所有的项为零，因此总和也为零，同 $\chi(0)$ 的定义相符.

因为两边都具有周期 2π，于是恒等式（23a）对所有 χ 均适用.

至于展式的系数确是由第 617 页上的公式(17)所定义的傅立叶系数，这是容易验证的.

现在 $\phi(x)$ 的傅立叶展式就直接从 $\phi(x)=\chi(\pi-x)$ 得到：

$$\phi(x)=2\sum_{\nu=1}^{\infty}(-1)^{\nu+1}\frac{\sin\nu x}{\nu}$$

$$=2\left(\sin x-\frac{1}{2}\sin 2x+\frac{1}{3}\sin 3x-+\cdots\right). \quad (23b)$$

只要用条件 $|x|<\pi-\varepsilon$ 把点 x 与间断点 $x=\pm\pi$ 隔离开来，这里的收敛就是一致的.

对于 $x=\pi/2$，我们又得到莱布尼兹级数

$$\frac{\pi}{2}=2\left(1-\frac{1}{3}+\frac{1}{5}-+\cdots\right).$$

应当提到的是，χ 和 ϕ 这两个级数并不绝对收敛；当 $x=\frac{\pi}{2}$ 时，绝对值确实形成发散级数

$$2\sum_{1}^{\infty}\frac{1}{2\nu-1}.$$

作为以连续函数为其项的无穷级数的一个例，公式（23b）是值得注意的，这个无穷级数对所有的 x 收敛，但是以一个非连续函数作为它的和，即以分段线性函数 $\phi(x)$ 作为它的和. 级数的每一个部分和是连续的，因为任何有限多个连续函数的和必然还是连续的. 由于连续函数的一致收敛的无穷级数具有一个连续和，所以傅立叶级数不可能在 ϕ 的不连续点 x 的一个邻域内一致收敛，这种点是 $x=\pm\pi$, $\pm 3\pi\cdots$. 第 608 页的图 8.4 用图说明了逐次的部分和(它们是三角多项式并且是连续函数)怎样在一个连续的区间内一致地逼近分段线性函数 $\frac{1}{2}\varphi(x)$，但是越靠近端点时，这些函数变化越快.

e. 关于傅立叶展开的主要定理

傅立叶系数. 有了上述的准备之后，一大类函数的可展性是容易确定的. 对于以 2π 为周期的函数 $f(x)$ 这样的展式的形式是

$$f(x)=\frac{1}{2}a_0+\sum_{\nu=1}^{\infty}(a_\nu\cos\nu x+b_\nu\sin\nu x), \tag{24a}$$

或记成复的形式

$$f(x)=\sum_{\nu=-\infty}^{\infty}\alpha_\nu e^{i\nu x}. \tag{24b}$$

我们先假定对于函数 $f(x)$ 有一致收敛的展式(24a)或(24b). 于是我们能够在这两个展式中分别确定系数 a_ν, b_ν 和 α_ν, 只要在这两个展式中分别乘以 $\cos\mu x$, $\sin\mu x$ 和 $e^{-i\mu x}$, 并且利用正交关系(见第 292 页和第 612 页)

$$\int_{-\pi}^{\pi}\sin\nu x\sin\mu x dx=\int_{-\pi}^{\pi}\cos\nu x\cos\mu x dx$$

$$=\begin{cases}0, & \text{如果 } \mu\neq\nu,\\ \pi, & \text{如果 } \mu=\nu\neq 0,\end{cases}$$

$$\int_{-\pi}^{\pi}\sin\nu x\cos\mu x dx=0, \tag{25}$$

$$\int_{-\pi}^{\pi}e^{i\nu x}e^{-i\mu x}dx=\begin{cases}0, & \text{如果 } \mu\neq\nu,\\ 2\pi, & \text{如果 } \mu=\nu,\end{cases}$$

从 $-\pi$ 到 π 积分就可以了. 把积分变量写成 t, 我们立即得到公式

$$a_\mu=\frac{1}{\pi}\int_{-\pi}^{\pi}f(t)\cos\mu t dt,\quad b_\mu=\frac{1}{\pi}\int_{-\pi}^{\pi}f(t)\sin\mu t dt, \tag{26a}$$

这里 $\mu=0, 1, 2, \cdots$; 而

$$\alpha_\mu=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(t)e^{-i\mu t}dt, \tag{26b}$$

这里 $\mu=0, \pm 1, \pm 2, \cdots$.

于是，如果 $f(x)$ 能够真的展开成为一个一致收敛的级数(24a)或(24b), 则系数只能取由公式(26a)和(26b)所确定的值. 但

是，即使这个颇含尝试性的步骤没有验证为合理，这些公式 (26a) 或(26b)对在区间 $-\pi \leqslant x \leqslant \pi$ 上连续或分段连续的每一个函数 $f(x)$ 仍然能确定出数列 a_ν, b_ν 和 α_ν 来，它们称为傅立叶系数.

对于一个给定的函数 $f(x)$，我们用 (26a, b) 这样定义的系数来形成傅立叶部分和

$$s_n(x) = \frac{1}{2} a_0 + \sum_{\nu=1}^{n} (a_\nu \cos \nu x + b_\nu \sin \nu x)$$

或

$$s_n(x) = \sum_{\nu=-n}^{\nu=n} \alpha_\nu e^{i\nu}.$$

我们的任务是要证明，当 $n \to \infty$ 时这些傅立叶和确实是收敛的，并且极限就是函数 $f(x)$.

现在我们叙述

主要定理 用傅立叶系数 (26a) 或 (26b) 形成的傅立叶级数

$$\frac{1}{2} a_0 + \sum_{\nu=1}^{\infty} (a_\nu \cos \nu x + b_\nu \sin \nu x) \qquad (27a)$$

或

$$\sum_{\nu=-\infty}^{\infty} \alpha_\nu e^{i\nu x} \qquad (27b)$$

对于任一周期为 2π 的分段连续的，有分段连续的一阶和二阶导数[1]的函数 $f(x)$ 来说，它收敛到函数值 $f(x)$. 这里在间断点上，$f(x)$ 的值必须由下式来确定

$$f(x) = \frac{1}{2} [f(x+0) + f(x-0)]. \qquad (27c)$$

证明[2] 为了证明，我们把系数的积分表达式 (26a) 代入 n 阶

---

1) 我们再次指出，这个定理能够对广泛得多的函数类（见第 8.6 节中的例）来证明. 然而，这里所叙述的结果对大部分的应用是足够的.

2) 这里我们只给出对于 f 的表达式在级数 (27a) 的形式下的证明. 级数 (27b) 只需根据第 615 页里的方程 (13b) 所给出的代换就得到了.

"傅立叶多项式"

$$s_n(x) = \frac{1}{2} a_0 + \sum_{\nu=1}^{n} (a_\nu \cos \nu x + b_\nu \sin \nu x),$$

而后交换积分与求和的顺序，得到

$$s_n(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \left[\frac{1}{2} + \sum_{\nu=1}^{n} (\cos \nu t \cos \nu x + \sin \nu t \sin \nu x) \right] dt,$$

或者利用余弦的加法定理，写成

$$s_n(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \left[\frac{1}{2} + \sum_{\nu=1}^{n} \cos \nu (t - x) \right] dt.$$

因此根据第 615 页的求和公式(14)就有

$$s_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \frac{\sin \left[\left(n + \frac{1}{2} \right) (t - x) \right]}{\sin \frac{1}{2} (t - x)} dt. \tag{28}$$

最后，令 $\tau = t - x$，并回忆起周期性容许我们把积分区间移动一个量 x（见第 602 页），我们就得到

$$s_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x + \tau) \frac{\sin \left(n + \frac{1}{2} \right) \tau}{\sin \frac{1}{2} \tau} d\tau, \tag{28a}$$

这里 x 当然是固定的.

现在我们证明当 $n \to \infty$ 时 $s_n(x)$ 趋于 $f(x)$；或者说

$$\lim_{n \to \infty} s_n(x) = \lim_{n \to \infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x + t) \frac{\sin \left(n + \frac{1}{2} \right) t}{\sin \frac{1}{2} t} dt$$

$$= f(x). \tag{29}$$

由于对所有的 x，$f(x) = \frac{1}{2} [f(x + 0) + f(x - 0)]$，我们有 [见第 616 页公式(15)]

$$s_n(x)-f(x)=\frac{1}{\pi}\int_0^{\pi}\frac{[f(x+t)-f(x+0)]}{2\sin\frac{1}{2}t}\sin\left(n+\frac{1}{2}\right)tdt$$

$$+\frac{1}{\pi}\int_{-\pi}^{0}\frac{[f(x+t)-f(x-0)]}{2\sin\frac{1}{2}t}\sin\left(n+\frac{1}{2}\right)tdt.$$

如果我们现在能够证明,变量 t 的函数 $[f(x+t)-f(x+0)]/\left(2\sin\frac{1}{2}t\right)$ 和 $[f(x+t)-f(x-0)]\Big/\left(2\sin\frac{1}{2}t\right)$ 及其一阶导数分别在区间 $0\leqslant t\leqslant\pi$ 和 $-\pi\leqslant t\leqslant 0$ 上是分段连续的,则根据我们的基本引理(第 618 页),当 $n\to\infty$ 时,右边的两个积分同时趋于 0,从而立即得到公式(29).

这样,如果我们能说明,在 f, f', f'' 是分段连续的条件下,对于一个固定的 x,用等式

$$\phi(t)=\frac{f(x+t)-f(x+0)}{2\sin\frac{1}{2}t},\quad 0<t<\pi,$$

$$\phi(t)=\frac{f(x+t)-f(x-0)}{2\sin\frac{1}{2}t},\quad -\pi<t<0$$

定义的 t 的函数是分段连续的并有分段连续的一阶导数，则主要定理就证明了.

要证实对于商 $\phi(t)$ 这些条件是满足的,首先我们注意到仅在 $t=0$ 时分母为 0,因此,除去可能靠近 $t=0$ 外,ϕ 及其一阶导数是分段连续的. 只是在奇点 $t=0$ 有可能丧失可微性. 因此,我们必须做的全部工作就是证明，如果 t 分别从正的或负的值趋于 0 时,$\phi(t)$ 及其导数 $\phi'(t)$ 有极限.我们将证明,这些极限确实是存在的,并且它们分别具有值

$$\phi(+0)=f'(x+0),\qquad \phi(-0)=f'(x-0)$$

和

$$\phi'(+0)=\frac{1}{2}f''(x+0),\qquad \phi'(-0)=\frac{1}{2}f''(x-0).$$

为证明起见,我们由 $\phi(t)=g(t)h(t)$ 而引入函数 $g(t)$,这里因子 $h(t)$ 用下式定义

$$h(t)=\frac{t}{2\sin(t/2)},\qquad t\neq 0,\ h(0)=1.$$

对 $h(t)$ 来说,由于 $h(0)=1$, $h'(0)=0$, 因而在整个区间 $-\pi\leqslant t\leqslant\pi$ 内,我们有(见第五章第 490 页)一个连续且连续可微的函数 $h(t)$; 因此,当 $t\to 0$ 时, $g(t)$ 和 $\varphi(t)$ 的极限值以及 $g'(t)$ 和 $\varphi'(t)=gh'+g'h$ 的极限值相同.

现在在区间 $0<t<\pi$ 里(见第五章第 488 页关于不定表达式的一般评述),根据微分学的中值定理有

$$g(t)=\frac{f(x+t)-f(x+0)}{t}=f'(x+\xi),$$

中值 ξ 位于 0 和 t 之间;因此当 t 趋于 0, 因而 ξ 也趋于 0 时

$$g(+0)=f'(x+0).$$

对于导数我们又得到表达式

$$g'(t)=\frac{tf'(x+t)+f(x+0)-f(x+t)}{t^2},$$

其中分子和分母当 $t\to 0$ 时都趋于 0, 并且分别具有导数 $tf''(x+t)$ 和 $2t$. 要确定当 $t\to 0$ 时的极限,我们用广义的中值定理(参看第 236 页)求得

$$g'(t)=\frac{\eta f''(x+\eta)}{2\eta}=\frac{1}{2}f''(x+\eta),$$

η 界于 0 和 t 之间. 当 $t\to 0$ 时,我们有 $\eta\to 0$, 因此,如前面说过的,

$$g'(+0)=\phi'(+0)=\frac{1}{2}f''(x+0).$$

同样的推理适用于 t 的负值. 于是我们对引理的应用是合理的,因而建立了主要定理.

可以再次指出,我们所获得的结果对微积分学及其应用中出

现的一切需要是足够了．可是从狄里克莱的最初的工作出发的数学家们的理论上的兴趣，常常以更大的普遍性为目标，即企图去展开更广泛的一类函数[1]． 这些努力已经刺激了对函数和积分概念的更加精密的分析，并导致作为一个有吸引力的专门领域的高等傅立叶分析的发展，不过，这已超出本书的范围．

8.5 傅立叶级数的例

a. 预先说明

我们始终假定函数 $f(x)$ 的周期是 2π.

如果 $f(x)$ 是一个偶函数（参看第 30 页），则显然 $f(x)\sin\nu x$ 是奇函数而 $f(x)\cos\nu x$ 是偶函数，所以

$$b_\nu=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin\nu x dx=0,$$

因而我们得到一个“余弦级数”．另一方面，如果函数 $f(x)$ 是奇函数，则

$$a_\nu=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos\nu x dx=0,$$

因而我们得到一个“正弦级数”.[2]

b. 函数 $\phi(x)=x^2$ 的展开式

对于偶函数 x^2，用两次分部积分，我们有

1）可注意的是，存在连续函数不能展成傅立叶级数的例子． 此外，还存在这样的例子，函数 $f(x)$ 由一个收敛的三角级数来表示，但这三角级数却不是以式（26）为系数的傅立叶级数．这些例子表明，对于精密的研究而言，在一般的三角级数与特殊的傅立叶级数之间作出区分是必要的．比连续性还要更严的限制条件确实是合宜的、而且对于我们，它们还说明了这样一个事实，即使是在我们的主要定理里所假定的限制和在第 8.6 节中所给的推广中，比真正需要的要严格得多（见 A. Zygmund 的 «Trigonometrical Series» 的一般理论，1952 年 Chelsea 出版公司）．

2）因而，如果函数 $f(x)$ 最初只是在区间 $0<x<\pi$ 上给定的，则我们能够把它或者作为一个奇函数或者作为一个偶函数开拓到区间 $-\pi<x<0$ 上，从而对于较小的区间 $0<x<\pi$ 得到的不是正弦级数就是余弦级数．

$$a_\nu = \frac{2}{\pi}\int_0^\pi x^2\cos\nu x dx = (-1)^\nu\frac{4}{\nu^2} \quad (\nu > 0),$$

$$a_0 = \frac{2\pi^2}{3},$$

因而我们得到展开式

$$x^2 = \frac{\pi^2}{3} - 4\left(\frac{\cos x}{1^2} - \frac{\cos 2x}{2^2} + \frac{\cos 3x}{3^2} - + \cdots\right). \quad (30)$$

逐项微分这个级数并除以 2，我们就在形式上回到了第 623 页上的级数(23b)，这个级数先前曾对 $\phi(x) = x$ 得到过.

c. $x\cos x$的展开式

(见图 8.9)对于这个奇函数,我们有

$$a_\nu = 0, \qquad b_\nu = \frac{2}{\pi}\int_0^\pi x\cos x\sin\nu x dx.$$

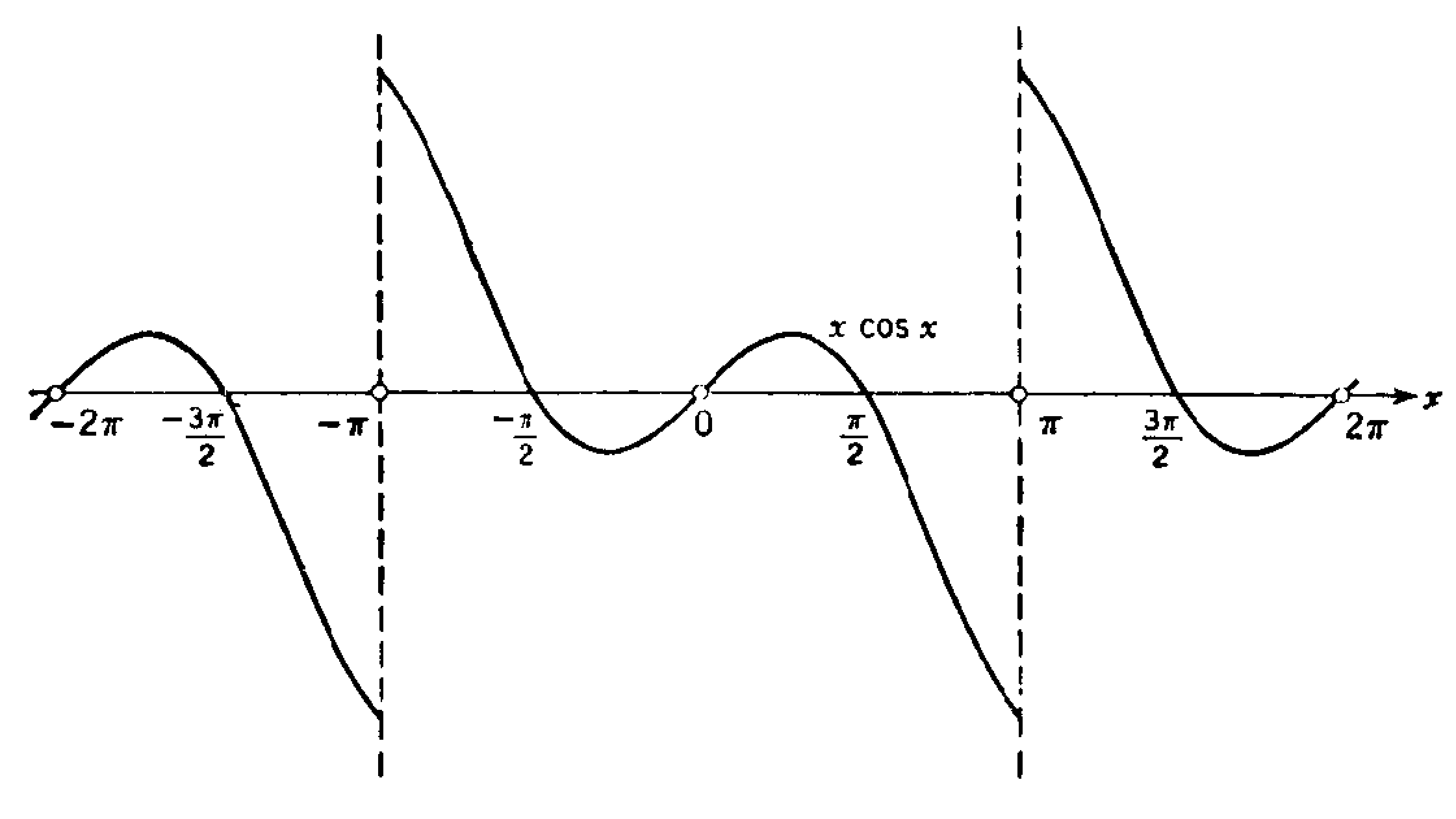


图 8.9

利用公式

$$\int_0^\pi x\sin\mu x dx = (-1)^{\mu+1}\frac{\pi}{\mu} \quad (\mu = 1, 2, \cdots),$$

我们求得

$$b_\nu = \frac{2}{\pi}\int_0^\pi x\cos x\sin\nu x dx$$

$$= \frac{1}{\pi}\int_0^{\pi} x[\sin(\nu+1)x + \sin(\nu-1)x]dx$$

$$= (-1)^{\nu}\left(\frac{2\nu}{\nu^2-1}\right) \qquad (\nu = 2, 3, \cdots),$$

$$b_1 = -\frac{1}{2}.$$

因此，我们得到级数

$$x\cos x = -\frac{1}{2}\sin x + 2\sum_{\nu=2}^{\infty}\frac{(-1)^{\nu}\nu}{\nu^2-1}\sin \nu x. \qquad (31)$$

加上第 623 页上对 $\phi(x) = x$ 找到的级数(23b)，就得到

$$x(1+\cos x) = \frac{3}{2}\sin x + 2\left(\frac{\sin 2x}{1\cdot 2\cdot 3} - \frac{\sin 3x}{2\cdot 3\cdot 4} + \frac{\sin 4x}{3\cdot 4\cdot 5} - + \cdots\right). \qquad (31a)$$

若将在区间 $-\pi < x < \pi$ 内等于 $x\cos x$ 的函数周期地开拓到这个区间之外，则会出现间断点(参看图 8.7)，这与在 8.4d 节中早已考虑过的函数 $\phi(x)$ 所展示的情况一样。另一方面，将函数 $x(1+\cos x)$ 周期性地开拓后，所得函数在区间的端点上保持连续，而且它的导数也保持连续，因为间断性由因子 $1+\cos x$ 消除了，$1+\cos x$ 随同其导数在端点上都为零. 这就说明了级数(31)对所有的 x 一致收敛，而从这级数同常数项级数 $\frac{1}{1^3}+\frac{1}{2^3}+\frac{1}{3^3}+\cdots$ 作比较来看则是明显的.

d. 函数 $f(x) = |x|$

对于这个偶函数 $b_\nu = 0$，并且 $a_\nu = \frac{2}{\pi}\int_0^{\pi} x\cos \nu x dx$；用分部积分，我们容易得到

$$\int_0^{\pi} x\cos \nu x dx = \frac{1}{\nu}x\sin \nu x\Big|_0^{\pi} - \frac{1}{\nu}\int_0^{\pi}\sin \nu x dx$$

$$=\begin{cases}0, & \text{如果 } \nu \text{ 是偶数且} \neq 0, \\ -\dfrac{2}{\nu^2}, & \text{如果 } \nu \text{ 是奇数.}\end{cases}$$

因而,

$$|x| = \frac{1}{2}\pi - \frac{4}{\pi}\left(\cos x + \frac{\cos 3x}{3^2} + \frac{\cos 5x}{5^2} + \cdots\right). \quad (32)$$

令 $x = 0$, 我们得到值得注意的公式

$$\frac{\pi^2}{8} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \cdots. \quad (32a)$$

e. 一个分段常数函数

由下列等式定义的函数

$$f(x) = \operatorname{sgn} x = \begin{cases}-1, & \text{当 } -\pi < x < 0, \\ 0, & \text{当 } x = 0, \\ +1, & \text{当 } 0 < x < \pi,\end{cases}$$

如同在第 32 页的图 1.22 上指出的, 它是奇函数. 因此 $a_\nu = 0$, 并且

$$b_\nu = \frac{2}{\pi}\int_0^\pi \sin \nu x dx = \begin{cases}0, & \text{如果 } \nu \text{ 是偶数,} \\ \dfrac{4}{\pi\nu}, & \text{如果 } \nu \text{ 是奇数,}\end{cases}$$

所以这个函数的傅立叶级数是

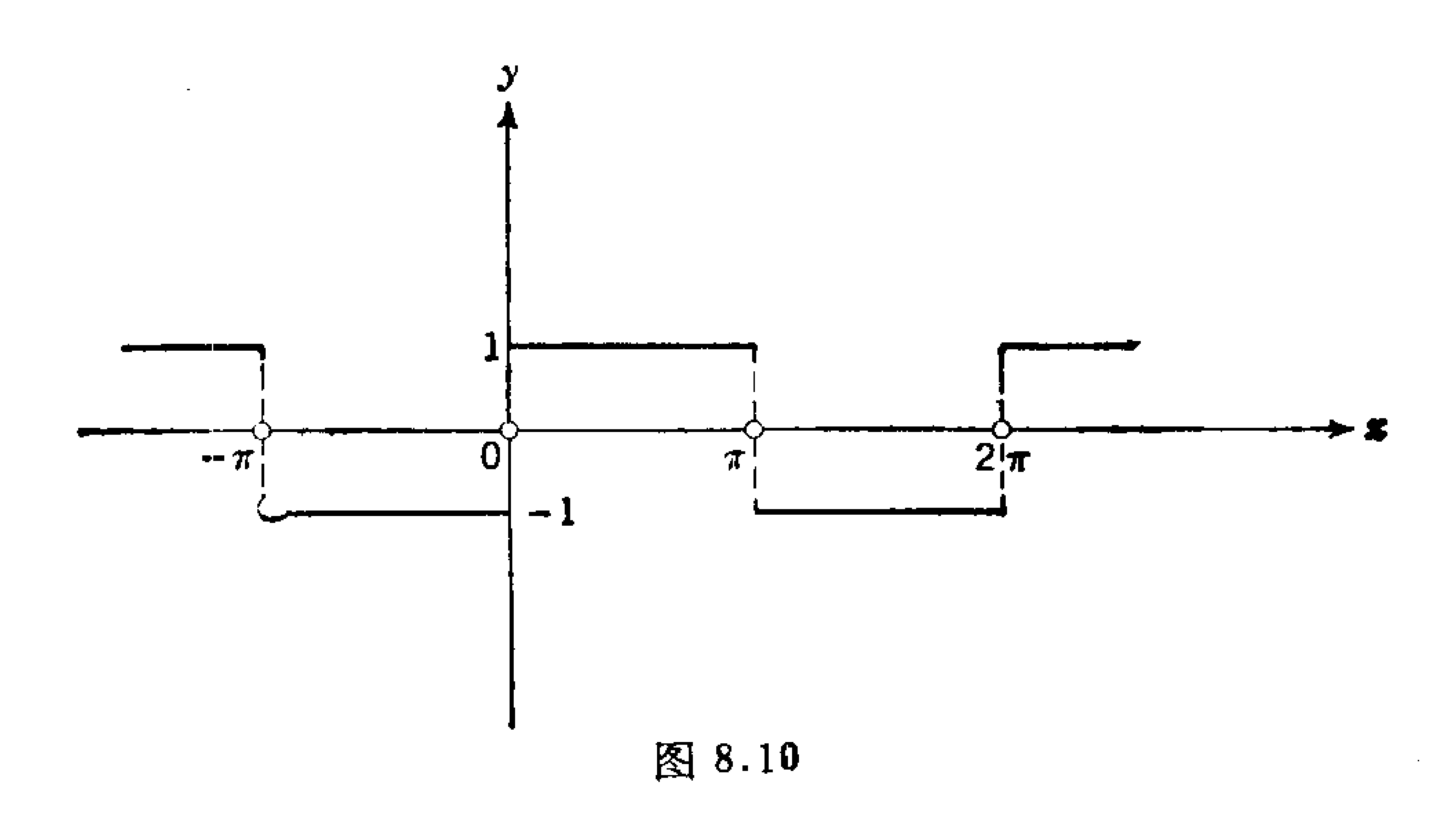


图 8.10

$$f(x) = \frac{4}{\pi}\left(\frac{\sin x}{1} + \frac{\sin 3x}{3} + \cdots\right). \tag{33}$$

特别地，当 $x = \frac{1}{2}\pi$，就再次得到莱布尼兹级数.

对在 (32) 中给出的 $|x|$，用逐项微分能够形式地导出级数 (33).

f. 函数 $|\sin x|$

偶函数 $f(x) = |\sin x|$ 能够展成余弦级数，其系数 a_ν 由下列计算给出：

$$\begin{aligned}
\frac{1}{2}\pi a_\nu &= \int_0^\pi \sin x \cos \nu x dx \\
&= \frac{1}{2}\int_0^\pi [\sin(\nu+1)x - \sin(\nu-1)x]dx \\
&= \begin{cases} 0, & \text{如果 } \nu \text{ 是奇数}, \\ \dfrac{-2}{\nu^2 - 1}, & \text{如果 } \nu \text{ 是偶数}. \end{cases}
\end{aligned}$$

于是我们用 2ν 代换 ν，得到

$$|\sin x| = \frac{2}{\pi} - \frac{4}{\pi}\sum_{\nu=1}^{\infty}\frac{\cos 2\nu x}{4\nu^2 - 1}. \tag{34}$$

g. $\cos \mu x$ 的展开式. 余切分解为部分分式. 正弦的无穷乘积

函数 $f(x) = \cos \mu x \quad (-\pi < x < \pi)$ 是偶函数，这里 μ 不是一个整数；因此 $b_\nu = 0$，而

$$\begin{aligned}
\frac{1}{2}\pi a_\nu &= \int_0^\pi \cos \mu x \cos \nu x dx \\
&= \frac{1}{2}\int_0^\pi [\cos(\mu+\nu)x + \cos(\mu-\nu)x]dx \\
&= \frac{1}{2}\left[\frac{\sin(\mu+\nu)\pi}{\mu+\nu} + \frac{\sin(\mu-\nu)\pi}{\mu-\nu}\right]
\end{aligned}$$

$$= \frac{\mu(-1)^{\nu}}{\mu^2-\nu^2}\sin\mu\pi.$$

于是我们有

$$\cos\mu x = \frac{2\mu\sin\mu\pi}{\pi}\left(\frac{1}{2\mu^2} - \frac{\cos x}{\mu^2-1^2} + \frac{\cos 2x}{\mu^2-2^2} + - \cdots\right). \tag{35}$$

这个从区间 $-\pi < x < \pi$ 周期性地开拓而成的以 2π 为周期的函数，在端点 $x = \pm\pi$ 保持连续. 令 $x = \pi$, 用 $\sin\mu\pi$ 除等式的两边并将 μ 写成 x, 我们得到等式

$$\cot\pi x = \frac{2x}{\pi}\left(\frac{1}{2x^2} + \frac{1}{x^2-1^2} + \frac{1}{x^2-2^2} + \cdots\right). \tag{36}$$

这就是余切化为部分分式的分解式（类似于第三章第 305 页讨论的有理函数的有限部分分式的分解式），它是数学分析的一个很重要的公式.

我们将上面的级数写成如下形式，

$$\cot\pi x - \frac{1}{\pi x} = -\frac{2x}{\pi}\left[\frac{1}{1^2-x^2} + \frac{1}{2^2-x^2} + \cdot\ \cdot\right].$$

如果 x 属于区间 $0 \leqslant x \leqslant q < 1$, 则右边第 n 项的绝对值小于 $2/[\pi(n^2-q^2)]$. 因此，级数在这个区间内一致收敛且能逐项积分. 两边乘以 π 并且积分，在左边我们得到

$$\pi\int_0^x\left(\cot\pi t - \frac{1}{\pi t}\right)dt = \log\frac{\sin\pi x}{\pi x} - \lim_{a\to 0}\log\frac{\sin\pi a}{\pi a}$$

$$= \log\frac{\sin\pi x}{\pi x},$$

在右边得到

$$\log\left(1-\frac{x^2}{1^2}\right) + \log\left(1-\frac{x^2}{2^2}\right) + \cdots = \lim_{n\to\infty}\sum_{\nu=1}^{n}\log\left(1-\frac{x^2}{\nu^2}\right);$$

于是

$$\log\frac{\sin\pi x}{\pi x} = \lim_{n\to\infty}\sum_{\nu=1}^{n}\log\left(1-\frac{x^2}{\nu^2}\right)$$

$$= \lim_{n\to\infty} \log \prod_{\nu=1}^{n} \left(1 - \frac{x^2}{\nu^2}\right)$$

$$= \log \lim \prod_{\nu=1}^{n} \left(1 - \frac{x^2}{\nu^2}\right).$$

如果我们从对数函数化到指数函数，那末我们有

$$\sin \pi x = \pi x\left(1 - \frac{x^2}{1^2}\right)\left(1 - \frac{x^2}{2^2}\right)\left(1 - \frac{x^3}{3^2}\right)\cdots. \qquad (36a)$$

这样我们就得到了关于正弦的无穷乘积的著名的表达式[1].

在这个结果中，令 $x = \frac{1}{2}$，我们就得到瓦里斯乘积

$$\frac{1}{2}\pi = \prod_{\nu=1}^{\infty} \frac{2\nu}{2\nu - 1} \cdot \frac{2\nu}{2\nu + 1} = \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5}\cdots.$$

这个公式在前面第 299 页上已经推导过.

h. 进一步的例

根据类似于前面的简短的计算，我们得到展开式的进一步的例.

按等式 $f(x) = \sin \mu x$ $(-\pi < x < \pi)$ 定义的函数 $f(x)$ 能展开为级数

$$\sin \mu x = -\frac{2 \sin \mu\pi}{\pi}\left(\frac{\sin x}{\mu^2 - 1^2} - \frac{2 \sin 2x}{\mu^2 - 2^2} + \frac{3 \sin 3x}{\mu^2 - 3^2} - + \cdots\right). \qquad (37)$$

令 $x = \frac{1}{2}\pi$ 并且使用关系式 $\sin \mu\pi = 2 \sin \frac{1}{2}\mu\pi \cos \frac{1}{2}\mu\pi$ 就得到余割的分解式，即得到函数 $1 \Big/ \cos \frac{1}{2}\mu\pi$ 化为部分分式的分解式；这个展式是

1) 这个公式特别有趣，因为它直接显示出在点 $x = 0, \pm 1, \pm 2, \cdots$ 函数 $\sin \pi x$ 为零. 在这点上它相应于当知道多项式的零点时该多项式的因式分解.

$$\pi \sec \pi x = \frac{\pi}{\cos \pi x} = 4\sum_{\nu=1}^{\infty} \frac{(-1)^{\nu}(2\nu-1)}{4x^2-(2\nu-1)^2},$$

这里我们已经将 $\frac{1}{2}\mu$ 改写成了 x.

类似于(35)和(37)，双曲函数 $\cosh \mu x$ 和 $\sinh \mu x$ $(-\pi < x < \pi)$ 的级数是

$$\cosh \mu x = \frac{2\mu}{\pi}\sinh \mu\pi\left(\frac{1}{2\mu^2} - \frac{\cos x}{\mu^2+1^2} + \frac{\cos 2x}{\mu^2+2^2} - \frac{\cos 3x}{\mu^2+3^2} + - \cdots\right),$$

$$\sinh \mu x = \frac{2}{\pi}\sinh \mu\pi\left(\frac{\sin x}{\mu^2+1^2} - \frac{2\sin 2x}{\mu^2+2^2} + \frac{3\sin 3x}{\mu^2+3^2} - + \cdots\right).$$

8.6 收敛性的进一步讨论

a. 结果

对傅立叶系数 a_ν, b_ν 的更细致的研究容易引出第 625 页 8.4e 节的主要定理的下列推论.

(a) 对于所有的周期函数在较弱的条件下,即只是 $f(x)$ 和它的一阶导数 $f'(x)$ 是分段连续的，或者如我们所说,函数是分段光滑的,则第 625 页上的傅立叶级数 (27a) 收敛到 $f(x)$.

(b) 如果分段光滑的周期函数 $f(x)$ 是连续的,则收敛是绝对和一致的.

(c) 如果分段光滑的函数 $f(x)$ 容许有跳跃间断点,则在不包含间断点的每一个闭区间上,收敛是一致的.

(b) 的证明依赖于一个简单的贝塞耳不等式,而(a) 和 (c)的证明将用到第 621 页 8.4 d 节的结果.

b. 贝塞耳不等式

这个不等式给出任何分段连续而不必可微的函数的傅立叶系

数的界限:

$$\frac{1}{2}a_0^2+\sum_{\nu=1}^{n}(a_\nu^2+b_\nu^2)\leqslant M^2. \tag{38}$$

这里,上界 $M^2=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)^2dx$ 是由函数 $f(x)$ 确定的一个数,它既不依赖于个别的傅立叶系数 a_ν, b_ν, 也不依赖于数 n. 对于复的傅立叶系数 α_ν [见第 615 页 (13a)], 贝塞耳 (Bassel) 不等式可立刻写成下述形式:

$$\sum_{\nu=-n}^{n}|\alpha_\nu|^2\leqslant\frac{1}{2\pi}\int_{-\pi}^{\pi}f(x)^2dx=\frac{1}{2}M^2. \tag{38a}$$

这一不等式是下述显然事实的直接推论:

$$\frac{1}{\pi}\int_{-\pi}^{\pi}\left[f(x)-\frac{1}{2}a_0-\sum_{\nu=1}^{n}(a_\nu\cos\nu x+b\nu\sin\nu x)\right]^2dx\geqslant 0.$$

在积分号下将平方展开,并注意到第 624 页的正交关系 (25) 以及第 617 页上傅立叶系数的定义(17),逐项积分,我们立刻得到上述形如 (38) 的贝塞耳不等式.

由于贝塞耳不等式的左边随着 n 而单调上升,而上界 M^2 是固定的,所以我们能令 $n\to\infty$ 通过取极限推出不等式

$$2\sum_{\nu=-\infty}^{\infty}|\alpha_\nu|^2=\frac{1}{2}a_0^2+\sum_{\nu=1}^{\infty}(a_\nu^2+b_\nu^2)\leqslant M^2 \tag{39}$$

是成立的. 对于分段连续的函数 $f(x)$ 的傅立叶系数,甚至当 f 不能用级数 (27a) 或 (27b) 表示时,这个不等式 (39) 仍适用.

附带指出,我们将在 8.7d 节中说明, 如果用等号代替不等号, 则贝塞耳不等式 (39) 仍然是成立的.

*c. 推论 (a), (b) 和 (c) 的证明

假定 $f(x)$ 本身是连续的,我们把贝塞耳不等式应用到它的分段连续的导数 $g(x)=f'(x)$ 上,该导数有傅立叶系数 $c_\nu=+\nu b_\nu$, $d\nu=-\nu a_\nu$, 因为我们可用分部积分立即求得这个结果 (因为积

出的项消去了）

$$c_\nu = \frac{1}{\pi}\int_{-\pi}^{\pi} f'(x)\cos \nu x dx$$

$$= + \frac{1}{\pi}\int_{-\pi}^{\pi} \nu f(x)\sin \nu x dx = + \nu b_\nu,$$

同理可求得 d_ν. [这里我们利用了 $f(x)$ 的连续性和周期性.] 于是我们有

$$\sum_{\nu=1}^{n} \nu^2(a_\nu^2 + b_\nu^2) = \sum_{\nu=1}^{n} (c_\nu^2 + d_\nu^2)$$

$$\leqslant \frac{1}{\pi}\int_{-\pi}^{\pi} g(x)^2 dx$$

$$= \frac{1}{\pi}\int_{-\pi}^{\pi} f'(x)^2 dx = M^2.$$

这个结果使我们能对 $f(x)$ 的傅立叶级数建立起具有正常数项的控制级数，该级数依照第 562 页所述保证了 (b) 中所述的绝对和一致的收敛性. 确实，根据柯西-施瓦兹不等式，对 ν 阶调和振动(参看第 16 页)我们首先有

$$|a_\nu \cos \nu x + b_\nu \sin \nu x|^2 \leqslant (a_\nu^2 + b_\nu^2)(\cos^2 \nu x + \sin^2 \nu x)$$

$$= a_\nu^2 + b_\nu^2;$$

然后对 $p = 1/\nu$, $q = \nu\sqrt{a_\nu^2 + b_\nu^2}$，利用不等式

$$pq \leqslant \frac{1}{2}(p^2 + q^2),$$

则对所有的 ν 我们有

$$|a_\nu \cos \nu x + b_\nu \sin \nu x| \leqslant \frac{1}{\nu}\nu\sqrt{a_\nu^2 + b_\nu^2}$$

$$\leqslant \frac{1}{2}\left[\frac{1}{\nu^2} + \nu^2(a_\nu^2 + b_\nu^2)\right].$$

由于最后的表达式关于 ν 的和是收敛的，所以我们已经构成了一个控制级数. 因此，傅立叶级数

$$\frac{1}{2}a_0+\sum_{\nu=1}^{\infty}(a_\nu\cos\nu x+b_\nu\sin\nu x)$$

一致收敛．于是它有一个和数 $s(x)$，$s(x)$ 是 x 的一个连续函数．为了证明确实有 $s(x)=f(x)$，我们采用一个技巧，考虑积分出来的函数

$$F(x)=\int_{-\pi}^{x}\left(f(t)-\frac{1}{2}a_0\right)dt.$$

显然，$F(x)$ 在 $-\pi\leqslant x\leqslant\pi$ 上是连续的；而且，在 $x=-\pi$ 和 $x=\pi$，F 有相同的值，因为

$$F(\pi)=\int_{-\pi}^{+\pi}f(t)dt-\pi a_0=0=F(-\pi).$$

从而 F 的周期开拓是连续的．又由于 F 的一阶和二阶导数都是分段连续的，所以函数 F 就表成了它的傅立叶级数．根据同样的推理(和前面一样以分部积分为基础)，当 $\nu\neq 0$ 时，F 的傅立叶系数是 $-(1/\nu)b_\nu$ 和 $(1/\nu)a_\nu$，所以

$$F(x)=\frac{1}{2}A_0+\sum_{\nu=1}^{\infty}\frac{1}{\nu}(-b_\nu\cos\nu x+a_\nu\sin\nu x),$$

式中含有某个常系数 A_0．现在按形式逐项微分而得到的级数已经知道是一致收敛的．于是按形式逐项微分是合法的(见第566页)，并且我们得到所希望的关系式

$$F'(x)=f(x)-\frac{1}{2}a_0=\sum_{\nu=1}^{\infty}(a_\nu\cos\nu x+b_\nu\sin\nu x).$$

要证明剩下的论述对于分段连续的并具有分段连续的导数 f' 的周期函数 f 成立，根据前面的结果，我们回忆起这些论述对于8.4d节的周期函数 $\chi(x)$ 是正确的，从而对函数 $\chi(x-\xi)$ 也是正确的，这个函数在点 ξ 有一个跳跃 2π．现在，如果函数 $f(x)$ 在点 $\xi_1,\xi_2,\cdots,\xi_m$ 上分别有跳跃 $\beta_1,\beta_2,\cdots,\beta_m$，则

$$f^*(x)=f(x)-\frac{1}{2\pi}\sum_{i=1}^{m}\beta_i\chi(x-\xi_i)$$

满足 (b) 的条件，因此具有一致收敛的傅立叶级数，于是对 $f(x)$ 证明了推论 (a) 和 (c).

d. 傅立叶系数的量阶. 傅立叶级数的微分法

前面收敛性的讨论说明了一个普遍事实：$f(x)$ 越光滑，也就是，周期函数 $f(x)$ 越多的高阶的导数连续的时候，当 $n \to \infty$ 时，傅立叶系数 a_ν, b_ν 就越快地收敛到零. 相应地，如果函数越光滑，傅立叶级数就收敛得越好. 我们确切地叙述如下：如果周期函数 $f(x)$ 具有直到 k 阶的连续的导数，并具有 $k+1$ 阶分段连续的导数，那末就存在一个只依赖于 $f(x)$ 和 k 的上界 B，使得

$$|a_\nu|,\ |b_\nu| < \frac{B}{\nu^{k+1}}. \tag{40}$$

如果我们利用分部积分法，证明几乎又是直接的 (见上面). 为简明起见，我们写成复数形式

$$a_\nu - ib_\nu = 2\alpha_\nu,$$

然后不断用分部积分法积分，直到在被积函数中出现因子 $f^{(k+1)}(x)$ 为止. 由于 $f(x)$, $f'(x)$, 等等的周期性和连续性，所以边界项彼此抵消了，因而

$$2\pi\alpha_\nu = \int_{-\pi}^{\pi} f(x)e^{-i\nu x}\,dx = -\frac{i}{\nu}\int_{-\pi}^{\pi} f'(x)e^{-i\nu x}dx$$
$$= \cdots = \left(\frac{-i}{\nu}\right)^{k+1}\int_{-\pi}^{\pi} f^{(k+1)}(x)e^{-i\nu x}dx.$$

因此，如果 $\frac{1}{2}B$ 是 $|f^{(k+1)}(x)|$ 的一个上界，则 $|\alpha_\nu| \leqslant \frac{1}{2}B/\nu^{k+1}$，由此就可推出不等式 (40).

更进一步值得注意的结果是，当 $k>2$ 时，傅立叶级数可逐项微分 $k-1$ 次，从而得到微分过的函数的傅立叶级数. 为了证明这一事实，我们注意到所有这些微分过的级数都以 $B\sum_{\nu=1}^{\infty}\frac{1}{\nu^2}$ 为收敛的控制级数，从而它们本身是绝对且一致收敛的(参看第七章第 562 页上的准则).

*8.7 三角多项式和有理多项式的近似法

a. 关于函数表示法的一般说明

要使这种“明显的表达式”表示函数成为可能需要对函数概念作什么样的限制，自从微积分早期以来就是一个挑战性的问题．函数常常不是从分析上给出的，而是从几何上或机械作图上，或由它们图形的几何描述给出的，这就可能在不同的区间内有不同的特性．

在十九世纪初期，傅立叶级数的发现是回答这个古老的问题的最光辉的一步；它揭示了“任意”函数确实能够用收敛的傅立叶级数来表示，而的确不必限于“解析”函数．而且，甚至傅立叶级数也并不包括全部连续函数：正如我们只提过一下而未加证明的那样，人们能够定义一些连续函数，其傅立叶级数由傅立叶系数形成，但并不收敛．

尤其值得注意的是，放弃在无穷级数中只是用增加高阶项的办法来达到其近似的原则，我们就可以对任何连续函数 $f(x)$ 构造 n 阶近似的三角多项式或有理多项式 $P_n(x)$，使其当 $n\to\infty$ 时在一闭区间上一致收敛到给定的函数 $f(x)$．

b. 外尔斯特拉斯逼近定理

我们来证明下列是诸密切相关定理．

(a) 如果 $f(x)$ 是闭区间 I 上的连续函数，I 包含在较大的区间 $-\pi<x<\pi$ 里面，则在 I 上能用一个周期为 2π 且阶 n 足够高的三角多项式来一致逼近 f．

(b) 在闭区间 I 上的任何连续函数 $f(x)$ 能够用 I 上的 x 的多项式 $P(x)$ 来一致地逼近．这一陈述是属于外尔斯特拉斯的，可用下面的推论来补充(见第 566 页)：

*(c) 如果 $f(x)$ 在 I 内有连续的导数，则近似多项式可以这样选择，使得多项式的导数 $P_n'(x)$ 一致地逼近导数 $f'(x)$．

(a) 的证明是相当直接的．首先我们用一个分段线性函数来

逼近 $f(x)$，它的图形是一个画在 $f(x)$ 图形中的多边形 $L_n(x)$（见图 8.11）．如果多边形的纵坐标是取在等间隔点 $x_1, x_2, \cdots, x_n$ 上，并且常数 $h=x_{\nu+1}-x_\nu$ 选择得足够小，则由于连续函数在区间 I 上一致连续（见第 106 页），显然，$L_n(x)$ 同 $f(x)$ 之差的绝对值可以小于任意选定的上界 $\varepsilon/2$．

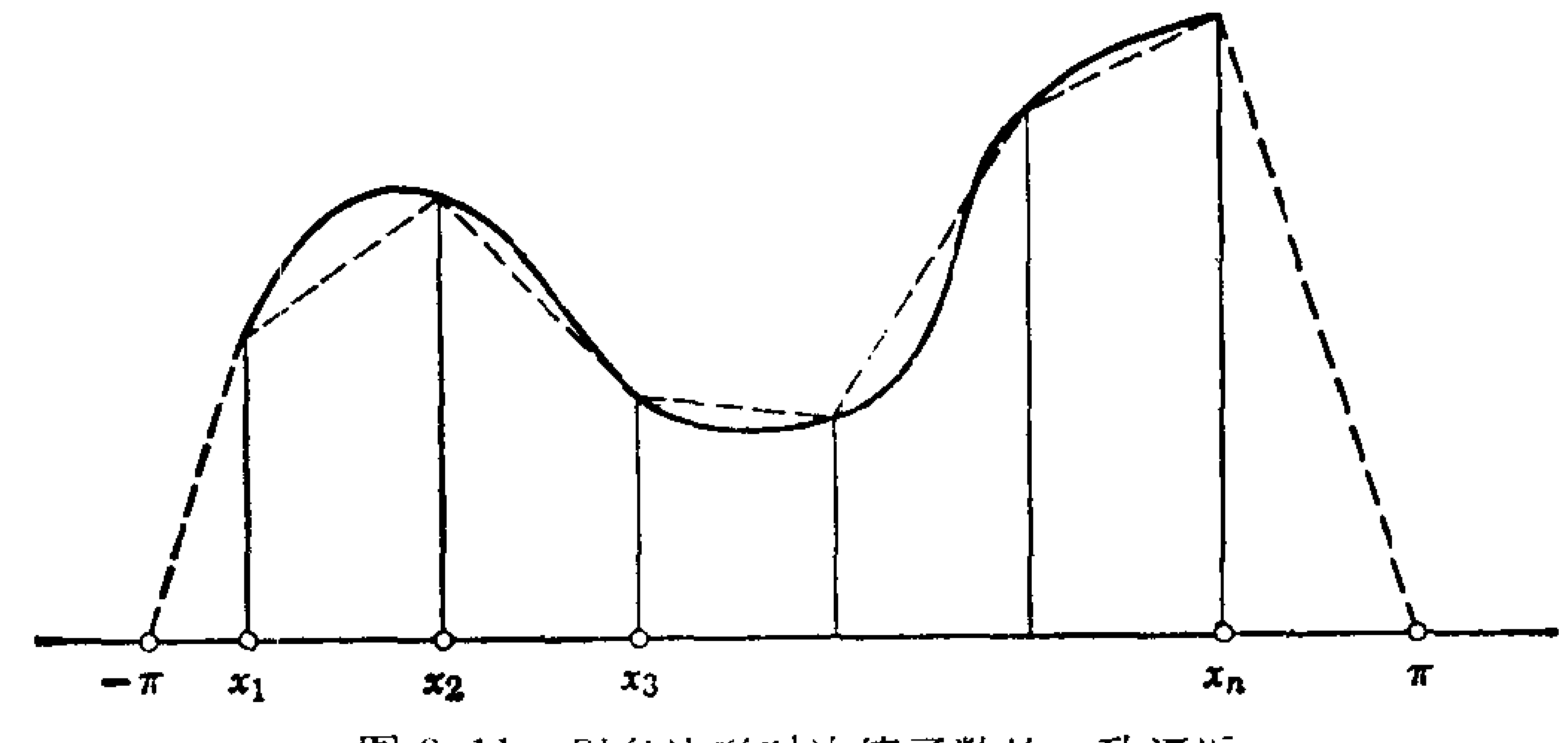


图 8.11　用多边形对连续函数的一致逼近

下一步（如图所示）用直线联结较大区间的端点 $-\pi$ 和 π，因而在闭区间 $-\pi \leqslant x \leqslant \pi$ 上把 $L_n(x)$ 开拓成一分段线性函数，仍称为 $L_n(x)$；在两端点上都为零的这个函数，现在能周期地开拓出去，并且按 8.6a 节可把它展成一致收敛的傅立叶级数，如果 m 充分大，该级数的多项式部分 $S_m(x)$ 同 $L_m(x)$ 之差的绝对值就小于 $\frac{\varepsilon}{2}$．现在 $|S_m-f| \leqslant |S_m-L_n|+|L_n-f|<\varepsilon$，因而(a)得证．[1)]

为了证明(b)，按照第 478 页 5.5b 节，在有限和 $S_m(x)$ 的每一项中，我们用带有一致小的余项的泰勒多项式来代替三角函数 $\cos \nu x$ 和 $\sin \nu x$；因此，组合上面这些近似式，我们构成一个多项式 $P_N(x)$，使得 $|P_N(x)-S_m(x)|<\frac{\varepsilon}{2}$，这里必须选择足够大的 N 以获得准确度 $\varepsilon/2$．综合起来，如果选择的 m 使 $|S_m(x)-$

1) 如果我们假定 $f(\pi)=f(-\pi)$，那么当 I 是整个区间 $-\pi \leqslant x \leqslant +\pi$ 时，这些结果仍然有效．在这里我们像以前一样选择一个近似的多边形 $L_n(x)$，只需使 $L_n(-\pi)=L_n(\pi)=f(-\pi)=f(\pi)$．

$f(x)| < \frac{\varepsilon}{2}$，我们在较小的区间上一定有 $|P_N(x)-f(x)| < \varepsilon$.

*c. 按算术平均值的傅立叶多项式的费叶尔三角近似式

有一个直接而又相当明确的构造近似多项式的方法能够非常简单地证明 8.7b 节的定理 (a). 这种构造是由下述值得注意的费叶尔定理提供的.

定理. 如果 $S_n(x)$ 是周期连续函数 $f(x)$ 的 n 阶傅立叶多项式，则当 $n \to \infty$ 时，算术平均值

$$F_n(x) = \frac{S_0(x) + \cdots + S_n(x)}{n+1}$$

一致收敛到 $f(x)$.

在通常的傅立叶逼近中，不管发生怎样的扰动振荡，这个定理由于取平均而保证了级数的收敛性.

证明. 这个定理的证明类似于傅立叶展式的主要定理的证明，然而它更简单些，因为这里振荡核 $\frac{\sin\left(n+\frac{1}{2}\right)x}{2\sin\frac{1}{2}x}$ 由正的"费叶尔核" $S_n(t) = \left(\frac{\sin\frac{1}{2}(n+1)t}{2\sin\frac{1}{2}t}\right)^2 \cdot \frac{2}{n+1}$ 来代替. 我们首先注意到，第 615 页上的函数 $\sigma_n(\alpha) = \frac{1}{2} + \cos\alpha + \cdots + \cos n\alpha$ 利用余弦的加法公式能写成下述形式

$$\sigma_n(\alpha) = \frac{\sin\left(n+\frac{1}{2}\right)\alpha}{2\sin\frac{1}{2}\alpha} = \frac{\sin\frac{1}{2}\alpha\sin\left(n+\frac{1}{2}\right)\alpha}{2\sin^2\frac{1}{2}\alpha}$$

$$= \frac{1}{2}\cdot\frac{\cos n\alpha - \cos(n+1)\alpha}{1-\cos\alpha}.$$

因此，我们得到公式

$$\frac{\sigma_0(\alpha)+\sigma_1(\alpha)+\cdots+\sigma_n(\alpha)}{n+1}=\frac{1}{2(n+1)}\frac{1-\cos(n+1)\alpha}{1-\cos\alpha}$$
$$=\frac{1}{2(n+1)}\left(\frac{\sin[(n+1)\alpha/2]}{\sin(\alpha/2)}\right)^2$$
$$=S_n(\alpha).$$

因为按 $\sigma_n(\alpha)$ 的定义[见第 615 页(14)]

$$\frac{1}{\pi}\int_{-\pi}^{\pi}\sigma_k(\alpha)d\alpha=1,$$

从而得到

$$\frac{1}{\pi}\int_{-\pi}^{\pi}S_n(\alpha)d\alpha=1.$$

现在[见第 626 页(28a)]

$$S_n(x)=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x+t)\sigma_n(t)dt$$

因而

$$F_n(x)=\frac{1}{\pi(n+1)}\int_{-\pi}^{\pi}f(x+t)[\sigma_0(t)+\cdots+\sigma_n(t)]dt$$
$$=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x+t)S_n(t)dt.$$

对任何正的 δ

$$f(x)-F_n(x)=\frac{1}{\pi}\int_{-\pi}^{\pi}[f(x)-f(x+t)]S_n(t)dt$$
$$=\frac{1}{\pi}\int_{-\delta}^{\delta}[f(x)-f(x+t)]S_n(t)dt$$
$$+\frac{1}{\pi}\int_{\delta}^{\pi}[f(x)-f(x+t)]S_n(t)dt$$
$$+\frac{1}{\pi}\int_{-\pi}^{-\delta}[f(x)-f(x+t)]S_n(t)dt.$$

现在因为 $f(x)$ 是一致连续的，因而我们能够选择一个 δ，使得对于所有在 $[-\pi,\pi]$ 上的 x 和对于 $|t|<\delta$ 有 $|f(x)-f(x+t)|$

$< \frac{1}{3}\varepsilon$. 此外 f 是有界的，譬如说 $|f| < M$. 因为按其定义

$$|S_n(t)| \leqslant \frac{1}{2(n+1)\sin^2(\delta/2)}, \quad 对\ \delta \leqslant |t| \leqslant \pi,$$

我们用 $S_n \geqslant 0$ 求出

$$|f(x) - F_n(x)| \leqslant \frac{\varepsilon}{3\pi}\int_{-\delta}^{\delta}|S_n(t)|dt$$

$$+ \frac{2\pi}{\pi}\int_{\delta}^{\pi}|S_n(t)|dt + \frac{2\pi}{\pi}\int_{-\pi}^{-\delta}|S_n(t)|dt$$

$$\leqslant \frac{\varepsilon}{3\pi}\int_{-\pi}^{\pi}S_n(t)dt + \frac{2\pi}{\pi}\frac{2\pi}{2(n+1)\sin^2(\delta/2)}$$

$$= \frac{\varepsilon}{3} + \frac{2\pi}{(n+1)\sin^2(\delta/2)}.$$

显然，当 n 足够大时，

$$|f(x) - F_n(x)| \leqslant \varepsilon,$$

因而定理得证.

*d. 在平均意义下的逼近和巴色瓦关系式

两个在闭区间 I 上连续的函数 $g(x)$ 和 $h(x)$ 的接近程度，从一致收敛的观点看，可用 $|g(x)-h(x)|$ 的最大值来度量. 在 I 上的连续函数 $\phi(x)$ 的最大绝对值称为它的**最大范数**，当 $n\to\infty$ 时，我们可以把函数序列 f_n 一致收敛到函数 f，说成差 $f-f_n$ 的最大范数趋于 0，或者也可以说成差 $f_n - f_m$ 的最大范数趋于 0.

很自然，对于傅立叶近似式而言(及本书范围之外的其它的重要数学理论)， 两个函数之间的偏差需给以另外的度量或 "范数(norm)"，或者考虑一个函数 $\phi(x)$ 离恒等于零的函数的距离是多少就够了. 这就是"二次平均"或者"**均方范数**" $\mu = \|\phi\|$，它由下述的平均值来定义

$$\mu^2 = \frac{1}{l}\int_I \phi(x)^2 dx = \|\phi\|^2,$$

式中 l 是区间 I 的长度. 这是一个比最大范数稍粗糙一些的度量,因为它的值小并不一定表示函数值处处都是小的.

作为一个例,函数 x^n 在区间 I: $0\leqslant x\leqslant 1$ 上的范数有值 $(2n+1)^{-\frac{1}{2}}$, 当 n 选得足够大时, 这个值可以任意小, 但是对于 $x=1$, 函数 x^n 等于 1.

如果当 $n\to\infty$ 时,二次范数 $\|f_n-f\|$ 趋于零,那末,我们就说在二次平均意义下 f_n 趋于 f.

由于下述的所谓三角不等式(它相当于数的三角不等式,见第 15 页)成立,所以把二次范数视为距离是颇为有用的. 这不等式即 $\|f+g\|\leqslant\|f\|+\|g\|$, 是对于两个函数 f 和 g 而言的, 它可由下面的推理立即得到: 取 $p=\dfrac{f(x)}{\|f\|}$, $q=\dfrac{g(x)}{\|g\|}$, 应用不等式 $pq\leqslant\dfrac{1}{2}(p^2+q^2)$, 并在 I 上积分,我们求得

$$\frac{1}{l}\int_I f(x)g(x)dx\leqslant\|f\|\cdot\|g\|.$$

现在

$$\begin{aligned}\|f+g\|^2&=\frac{1}{l}\int_I[f(x)+g(x)]^2dx\\&=\|f\|^2+\|g\|^2+\frac{1}{l}\int_I 2f(x)g(x)dx\\&\leqslant(\|f\|+\|g\|)^2\end{aligned}$$

或

$$\|f+g\|\leqslant\|f\|+\|g\|.$$

借助这些概念,我们可以把 8.6b 节的贝塞耳不等式看得更清楚. 我们首先指明: 在"二次平均"的意义下, 用一个其系数 c_ν, d_ν 是可以自由选择的 n 阶三角多项式

$$\begin{aligned}T_n&=\frac{c_0}{2}+\sum_{\nu=1}^{n}(c_\nu\cos\nu x+d_\nu\sin\nu x)\\&=\sum_{\nu=-n}^{n}\beta_\nu e^{i\nu x}\end{aligned}$$

$\left[\text{其中 } \beta_0 = \frac{c_0}{2},\ \beta_\nu + \beta_{-\nu} = c_\nu,\ i(\beta_\nu - \beta_{-\nu}) = d_\nu,\right]$ 去逼近一个给定的分段连续函数时，其最佳逼近是由傅立叶多项式

$$S_n = \frac{a_0}{2} + \sum_{\nu=1}^{n} (a_\nu \cos \nu x + b_\nu \sin \nu x)$$
$$= \sum_{\nu=-n}^{n} \alpha_\nu e^{i\nu x}$$

给出的，式中 a_ν, b_ν 和 α_ν 是实的和复的傅立叶系数，它们分别根据第 624 页公式 (26) 由 f 确定。

为了简明起见，证明用复的形式写。在区间 $I = [-\pi, \pi]$ 上对函数 $e^{i\nu x}$ 利用正交关系 (25) 容易得到：

$$\frac{1}{l}\int_I \left(f(x) - \sum_{\nu=-n}^{n} \beta_\nu e^{i\nu x}\right)^2 dx$$
$$= \frac{1}{l}\int_I \left[f(x)^2 - 2\sum_{\nu=-n}^{n} \beta_\nu f(x) e^{i\nu x} + \left(\sum_{\nu=-n}^{n} \beta_\nu e^{i\nu x}\right)^2\right] dx$$
$$= \|f\|^2 - 2\sum_{\nu=-n}^{n} \beta_\nu \alpha_{-\nu} + \sum_{\nu=-n}^{n} \beta_\nu \beta_{-\nu}$$
$$= \|f\|^2 - \sum_{\nu=-n}^{n} \alpha_\nu \alpha_{-\nu} + \sum_{\nu=-n}^{n} (\alpha_\nu - \beta_\nu)(\alpha_{-\nu} - \beta_{-\nu})$$
$$= \|f\|^2 - \sum_{\nu=-n}^{n} \alpha_\nu \bar{\alpha}_\nu + \sum_{\nu=-n}^{n} (\alpha_\nu - \beta_\nu)(\bar{\alpha}_\nu - \bar{\beta}_\nu)$$
$$= \|f\|^2 - \sum_{\nu=-n}^{n} |\alpha_\nu|^2 + \sum_{\nu=-n}^{n} |\alpha_\nu - \beta_\nu|^2;$$

显然，当把 β_ν 选为傅立叶系数 α_ν 时，即当 $\alpha_\nu = \beta_\nu$，或等价地，$c_\nu = a_\nu$, $d_\nu = b_\nu$ 时，上述表达式达到极小值。

利用上面得到的逼近结果，现在我们能够证明巴色瓦(Parseval)定理：

贝塞耳不等式

$$\frac{1}{2}a_0^2 + \sum_{\nu=1}^{n} (a_\nu^2 + b_\nu^2) \leqslant \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)^2 dx,$$

当 $n \to \infty$ 时，变为巴色瓦等式

$$\frac{1}{2} a_0^2 + \sum_{\nu=1}^{\infty} (a_\nu^2 + b_\nu^2) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)^2 dx,$$

适用于任何周期为 2π 的对一切 x 连续的函数.

证明. 根据外尔斯特拉斯关于三角多项式的逼近定理，我们可选择一个多项式序列 T_n 使得 $f(x) - T_n(x)$ 对 x 一致地趋于零. 于是，当 $n \to \infty$ 时也有

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} [f(x) - T_n(x)]^2 dx \to 0.$$

但是，按照我们上述的结果，在一切 n 阶傅立叶多项式中，在平均的意义下，傅立叶多项式

$$S_n(x) = \frac{a_0}{2} + \sum_{\nu=1}^{n} (a_\nu \cos \nu x + b_\nu \sin \nu x)$$

对 $f(x)$ 给出最佳的逼近，所以

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} [f(x) - S_n(x)]^2 dx \leqslant \frac{1}{2\pi} \int_{-\pi}^{\pi} [f(x) - T_n(x)]^2 dx.$$

从而得到

$$\lim_{n\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} [f(x) - S_n(x)]^2 dx = 0.$$

如同第 647 页一样，平方被积函数，我们就得到了巴色瓦关系式.

最后，我们注意到，如果 $f(x)$ 有几个跳跃间断点，巴色瓦关系式仍然保持有效. 这里我们省略了这个简单证明.

附　　录　　I

*A.I.1　周期区间的伸缩变换. 傅立叶积分定理

对我们的周期函数，可用任何区间 $-B \leqslant x \leqslant B$ 代换基本区间 $-\pi \leqslant x \leqslant \pi$. 用变换 $y = \pi x/B$ 将这个长度为 $2B$ 的区间转

换为区间 $-\pi \leqslant y \leqslant \pi$，并且将周期为 $2B$ 的函数 $f(x)$ 转换为周期为 2π 的函数 $g(y)=f(By/\pi)=f(x)$. 记为复的形式[见第 625 页公式(27b)]的主要定理，蕴含着

$$g(y)=\sum_{\nu=-\infty}^{\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi}g(t)e^{-i\nu(t-y)}dt.$$

根据这个变换，于是

$$f(x)=\frac{1}{2\pi}\sum_{\nu=-\infty}^{\infty}\frac{\pi}{B}\int_{-B}^{B}f(s)e^{-i\nu\pi(s-x)/B}ds, \tag{41}$$

这里积分变量用 $s=Bt/\pi$ 代替.

对于在区间 $-B\leqslant x\leqslant B$ 上每个分段光滑的函数，关系式(41)是有效的.

我们令 $\pi/B=h$, $\nu\pi/B=\nu h=u_\nu$，并且将(41)写成下述形式

$$f(x)=\frac{1}{2\pi}\sum_{\nu=-\infty}^{\infty}h\int_{-B}^{B}f(s)e^{-iu_\nu(s-x)}ds$$

$$=\frac{1}{2\pi}\sum_{\nu=-\infty}^{\infty}he^{iu_\nu x}H_\nu,$$

其中 $H_\nu=\int_{-B}^{B}e^{-iu_\nu s}f(s)ds$. 现在就 $B\to\infty$ 或 $\Delta u=h\to 0$ 取极限，这在形式上的可行性是明显的，因而得到

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{iux}du\int_{-\infty}^{\infty}e^{-ius}f(s)ds. \tag{42}$$

这就是傅立叶积分公式，这个公式将在第二卷中对一大类函数 f 给以严格证明. 这个公式可写成明显对称的形式，把它作为一个函数 $f(x)$ 和它的“傅立叶变换” $F(u)$ 之间的一对互反的积分关系式：

$$F(u)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}f(s)e^{-ius}ds, \tag{43}$$

$$f(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}F(u)e^{iux}dx. \tag{43a}$$

傅立叶积分公式(42)可写成不包含虚指数的形式. 我们只需利用表示式

$$e^{iux}e^{-ius} = e^{iu(x-s)} = \cos u(s-x) - i\sin u(s-x).$$

因为 $\sin u(s-x)$ 是 u 的奇函数,而 $\cos u(s-x)$ 是偶函数,从 $-\infty$ 到 $+\infty$ 对 u 的正弦项的积分为零,而积分余弦项得到从 0 到 ∞ 积分值的两倍. 因此

$$f(x) = \frac{1}{\pi}\int_0^{\infty} du \int_{-\infty}^{+\infty} f(s)\cos u(s-x)ds. \tag{43b}$$

*A.I.2 非连续点上的吉布斯现象

在一个跳跃间断点附近的傅立叶级数的收敛性质显示出一个值得注意的特点,这是吉布斯 (Gibbs) 在考查傅立叶多项式

$$S_n(x) = \frac{a_0}{2} + \sum_{\nu=1}^{n}(a_\nu \cos\nu x + b_\nu \sin\nu x)$$

的图形时发现的. 正如在第七章第 557 页中已经强调过的,在极限函数的一个间断点的附近,一个收敛序列的不一致收敛性能够用逼近函数的连续的图形不逼近于这个极限函数的不连续的图形的方法观察到.

在傅立叶展开式中,这些图形不是简单地逼近增补后的 $f(x)$ 的图形的,这里的增补是指,在跳跃位置 ξ 处,用垂直线段 $x=\xi$ 联结两个端点. 而是这样逼近的: S_n 的图形表示了波动,这波在接近 ξ 时,大约以总跳跃的9%超出纵坐标 $f(\xi+0)$ 和 $f(\xi-0)$ 的上下两边. 因此,由近似图形去逼近 $f(x)$ 的图形时,在 $x=\xi$,不仅要联结 $f(x)$ 图形上的两个点,而且要增加超过两个端点的垂直线段;见第 608 页和 609 页图 8.4 和 8.5.

从数学上分析这种情况是简单的,只需要讨论 8.4d 节的函数 $\chi(x)$ 的间断性就可以了,第 639 页上的一切跳跃间断都可化为 $\chi(x)$ 的情况.

对于正的 x,函数 $\frac{1}{2}\chi(x)$ [见第 622 页公式(23a)] 由下式

给出

$$\frac{1}{2}\chi(x)=\frac{1}{2}(\pi-x)=\sum_{\nu=1}^{\infty}\frac{\sin\nu x}{\nu},\quad 0<x<\pi.$$

根据第 615 页公式 (14) 的积分,我们得到

$$S_n(x)=\sum_{\nu=1}^{n}\frac{\sin\nu x}{\nu}=-\frac{1}{2}\chi+\int_0^x\frac{\sin\left(n+\frac{1}{2}\right)t}{2\sin\frac{1}{2}t}dt,$$

因此余项 $r_n(x)=\frac{1}{2}\chi(x)-S_n(x)$ 取形式

$$r_n(x)=\frac{1}{2}\pi-\int_0^x\frac{\sin\left(n+\frac{1}{2}\right)t}{t}dt+\rho_n(x),$$

式中

$$\rho_n(x)=\int_0^x\frac{2\sin\frac{1}{2}t-t}{2t\sin\frac{1}{2}t}\sin\left(n+\frac{1}{2}\right)tdt.$$

因为表达式 $\left(2\sin\frac{1}{2}t-t\right)\Big/2t\sin\frac{1}{2}t$ 是分段连续并且有分段连续的一阶导数，第 618 页上的引理就蕴含着对 $0<x<\pi$，当 $n\to\infty$ 时，$\rho_n(x)$ 一致地趋于0．此外，

$$\begin{aligned}\sigma_n(x)&=\frac{1}{2}\pi-\int_0^x\frac{\sin\left(n+\frac{1}{2}\right)t}{t}dt\\&=\frac{1}{2}\pi-\int_0^{(n+1/2)x}\frac{\sin t}{t}dt,\end{aligned}$$

当 $n\to\infty$ 时对于每一个各别正的 x 趋于 0 (见第 621 页)．但是收敛是不一致的．显然，对于 $k=1,2,3,\cdots$ 在点 $x_k=2k\pi/(2n+1)$ 上，$\sigma_n(x)$ 的导数为零．容易看出，更确切地，$\sigma_n(x)$ 在点 $x_1, x_3, x_5, \cdots$ 有最小值，在点 $x_2, x_4, x_6, \cdots$ 有最大值．此外,在最小值点 σ_n 的值形成一个递增的序列．于是，对于正的

x 值，$\sigma_n(x)$ 以下面这个值作为它的"绝对"最小值：

$$\sigma_n(x_1) = \frac{1}{2}\pi - \int_0^\pi \frac{\sin t}{t} dt$$

$$= \frac{1}{2}\pi - \int_0^\pi \left(1 - \frac{1}{3!}t^2 + \frac{1}{5!}t^4 - + \cdots\right) dt$$

$$= \pi\left(\frac{1}{2} - 1 + \frac{\pi^2}{2\cdot 3\cdot 3} - \frac{\pi^4}{2\cdot 3\cdot 4\cdot 5\cdot 5}\right.$$

$$\left. + \frac{\pi^6}{2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7\cdot 7} - + \cdots\right)$$

$$\approx -0.090\cdots\pi.$$

当 n 很大时，余项 r_n 近似地等于 σ_n。因此，当 n 很大时，近似多项式 S_n 超过函数 χ 大约 $(9/100)\pi$，即约为在原点处函数的右极限和左极限之差的 9%。因此，$S_n(x)$ 的图形的振荡分支确实超过 $\chi(x)$ 图形的高度并且呈现出上面描述的极限现象.

容易看出，和数 $S_n(x)$ 的费叶尔平均值是没有吉布斯现象的.

*A.I.3 傅立叶级数的积分

一般讲，如我们已看到的(第 563 页)，如果一个无穷级数是一致收敛的，它就能逐项积分. 然而，对于傅立叶级数我们有一个值得注意的结果，就是逐项积分总是可能的. 我们说：**如果 $f(x)$ 是一个在 $-\pi \leqslant x \leqslant \pi$ 上分段连续的函数，它具有形式上的傅立叶展式**

$$\frac{1}{2}a_0 + \sum_{\nu=1}^{\infty}(a_\nu \cos \nu x + b_\nu \sin \nu x),$$

则对任意两点 x_1, x_2 有

$$\int_{x_1}^{x_2} f(x)dx = \int_{x_1}^{x_2}\frac{1}{2}a_0 dx + \sum_{\nu=1}^{\infty}\int_{x_1}^{x_2}(a_\nu \cos \nu x + b_\nu \sin \nu x)dx,$$

就是说傅立叶级数能逐项积分. 此外，对于固定的 x_1，右边的级数对 x_2 一致收敛.

这个定理的值得注意的部分是，不仅不需要假定级数的一致收敛性，而且甚至也不需要利用它的收敛性。

为了证明定理，如同第 639 页一样定义

$$F(x)=\int_{-\pi}^{x}\left[f(t)-\frac{1}{2}a_0\right]dt.$$

$F(x)$ 是连续的并有分段连续的导数；此外，它满足条件 $F(\pi)=F(-\pi)=0$，于是把它周期地开拓之后它仍是连续的。因此，$F(x)$ 的傅立叶级数

$$\frac{1}{2}A_0+\sum_{\nu=1}^{\infty}(A_\nu\cos\nu x+B_\nu\sin\nu x)$$

一致收敛到 $F(x)$。利用分部积分法，当 $\nu\neq 0$，我们得到傅立叶系数的值

$$A_\nu=\frac{1}{\pi}\int_{-\pi}^{\pi}F(t)\cos\nu t dt=-\frac{1}{\pi}\int_{-\pi}^{\pi}f(t)\frac{\sin\nu t}{\nu}dt=-\frac{b_\nu}{\nu},$$

$$B_\nu=\frac{1}{\pi}\int_{-\pi}^{\pi}F(t)\sin\nu t dt=\frac{1}{\pi}\int_{-\pi}^{\pi}f(t)\frac{\cos\nu t}{\nu}dt=\frac{a_\nu}{\nu}.$$

于是级数

$$F(x_2)-F(x_1)=\sum_{\nu=1}^{\infty}[A_\nu(\cos\nu x_2-\cos\nu x_1)+B_\nu(\sin\nu x_2-\sin\nu x_1)]$$

$$=\sum_{\nu=1}^{\infty}\left[-\frac{b_\nu}{\nu}(\cos\nu x_2-\cos\nu x_1)+\frac{a_\nu}{\nu}(\sin\nu x_2-\sin\nu x_1)\right]$$

关于 x 一致收敛。把 $F(x)$ 换成 $\int_{\pi}^{x}\left[f(x)-\frac{1}{2}a_0\right]dx$，我们得到关系式

$$\int_{x_1}^{x_2}\left[f(x)-\frac{1}{2}a_0\right]dx=\sum_{\nu=1}^{\infty}\int_{x_1}^{x_2}(a_\nu\cos\nu x+b_\nu\sin\nu x)dx,$$

这正是所要证明的。

附　录　II

*A.II.1　伯努利多项式及其应用

a. 定义及傅立叶展式

在泰勒级数的推导过程中（第 474 页），以 ξ 为参数的 x 的多项式 $P_n(x)=(\chi-\xi)^n/n!$，$n\geqslant 1$ 起了作用. 这一序列多项式是这样刻划的：每一个多项式 P_{n+1} 是 P_n 的原函数即 $P'_{n+1}(x)=P_n(x)$，而且，$P_n(\xi)=0$ 和 $P_0(x)\equiv 1$.

我们现在按逐次积分来构造另一个重要的多项式序列，即伯努利多项式序列，然后我们将把这些多项式开拓为周期函数并展为傅立叶级数.

对 $0\leqslant x\leqslant 1$，伯努利多项式 $\phi_n(x)$ 递归地用下列关系式来定义：

$$\phi'_n(x)=\phi_{n-1}(x),\qquad \phi_0(x)=1, \tag{44a}$$

$$\int_0^1\phi_n(x)dx=0,\qquad \text{当 } n>0. \tag{44b}$$

对已知的 $\varphi_0,\varphi_1,\cdots,\varphi_{n-1}$，由条件 (44a) 来确定 φ_n，其中含有一个任意的积分常数，这个常数可由条件 (44b) 完全固定. 根据数学归纳法我们立即看到，ϕ_n 是一个具有有理系数的 n 阶多项式. 前几个伯努利多项式是容易计算的：

$$\phi_0(x)=1,$$

$$\phi_1(x)=x-\frac{1}{2},$$

$$\phi_2(x)=\frac{1}{2}x^2-\frac{1}{2}x+\frac{1}{12},$$

$$\phi_3(x)=\frac{1}{6}x^3-\frac{1}{4}x^2+\frac{1}{12}x,$$

$$\phi_4(x)=\frac{1}{24}x^4-\frac{1}{12}x^3+\frac{1}{24}x^2-\frac{1}{720}.$$

当 $n>1$，按 (44a, b) 我们有

$$\phi_n(1)-\phi_n(0)=\int_0^1 \phi_n'(t)dt=0.$$

于是，多项式函数 ϕ_n 可以从基本区间 $0\leqslant x\leqslant 1$ 开拓到所有的 x 使其成为以 1 为周期的连续的周期函数 $\psi_n(x)$，称为伯努利函数，而函数 $\psi_1(x)$ 同非连续函数 $\frac{1}{2\pi}\phi(2\pi x-\pi)$ 相重合，并且能够表示为傅立叶级数（见第 623 页公式 (23b)）

$$\psi_1(x)=-\frac{1}{\pi}\left(\frac{\sin 2\pi x}{1}+\frac{\sin 4\pi x}{2}+\frac{\sin 6\pi x}{3}+\cdots\right). \quad (45a)$$

通过逐次积分，我们就得到

$$\psi_n(t)=(-1)^{(n/2)+1}\cdot\frac{2}{(2\pi)^n}\sum_{k=1}^{\infty}\frac{\cos 2\pi kt}{k^n},$$

当 n 是偶数， (45b)

$$\psi_n(t)=(-1)^{(n+1)/2}\cdot\frac{2}{(2\pi)^n}\sum_{k=1}^{\infty}\frac{\sin 2\pi kt}{k^n},$$

当 n 是奇数. (45c)

在原来的区间 $0\leqslant x\leqslant 1$ 上，周期函数 $\psi_n(t)$ 恒等于伯努利多项式 $\phi_n(t)$.

当 n 是偶数时，ψ_n 是偶函数，当 n 是奇数时，ψ_n 是奇函数；等价地说

$$\psi_n(-x)=(-1)^n\psi_n(x). \quad (45d)$$

在逐个的伯努利多项式中，常数项组成一个值得注意的有理数序列

$$b_n=\phi_n(0)=\begin{cases}\psi_n(0), & \text{当}\quad n\neq 1,\\ -\frac{1}{2}, & \text{当}\quad n=1.\end{cases} \quad (46a)$$

从傅立叶展式我们立即得到

$$b_n=0,\quad \text{当}\quad n=3,5,\cdots,\text{为奇数时}, \quad (46b)$$

$$b_n = (-1)^{n/2+1}\frac{2}{(2\pi)^n}\sum_{k=1}^{\infty}\frac{1}{k^n},\ \text{当}\ n = 2, 4, \cdots,\ \text{为偶数时}. \tag{46c}$$

而且，显然当 $n = 2m$ 为偶数时，b_{2m} 的符号交替变换.

代替随 n 的增加而迅速减小的数 b_n，雅可布·伯努利引入下列多少更合适一点的数:

$$B_m = (-1)^{m-1}(2m)!\,b_{2m}, \tag{47}$$

我们称它为伯努利数.（数 $B^*_{2m} = (-1)^{m-1}B_m$ 与在第 590 页引入的伯努利数是恒等的，这在以后将成为明显的.）特别地，

$$B_1 = \frac{1}{6},\quad B_2 = \frac{1}{30},\quad B_3 = \frac{1}{42},\quad B_4 = \frac{1}{30},$$

$$B_5 = \frac{5}{66},\quad B_6 = \frac{691}{2730},\quad B_7 = \frac{7}{6},\cdots.$$

作为公式 (46c) 的一个结果，在

$$\sum_{k=1}^{\infty}\frac{1}{k^{2n}} = (-1)^{n-1}(2\pi)^{2n}\frac{1}{2}\,b_{2n} = \frac{(2\pi)^{2n}}{2(2n)!}B_n \tag{48}$$

里，当整数 $s = 2n$ 时，我们通过已知的这些数给出了黎曼 ζ 函数 $\zeta(s)$ 的明确的表达式(见第 587 页). 例如，我们得到值得注意的公式如

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots = \frac{\pi^2}{6} = \zeta(2)$$

及

$$1 + \frac{1}{2^4} + \frac{1}{3^4} + \frac{1}{4^4} + \cdots = \frac{\pi^4}{90} = \zeta(4).$$

当 $n \to \infty$ 时，数 b_n 和 B_n 分别趋于零和无穷大. 因为，首先，我们有

$$1 < \sum_{k=1}^{\infty}\frac{1}{k^{2n}} \leqslant \sum_{k=1}^{\infty}\frac{1}{k^2} = \frac{\pi^2}{6} < 2.$$

于是

$$2(2\pi)^{-2n} < |b_{2n}| < 4(2\pi)^{-2n}.$$

因为 $2\pi > 1$ 并且当 $n \to \infty$ 时 $(2\pi)^{-2n} \to 0$，我们有 $b_{2n} \to 0$，而 $b_{2n+1} = 0$. 此外，

$$B_n = (2n)!\,|b_{2n}| > 2(2n)!\,(2\pi)^{-2n};$$

容易看出，右边趋向无穷.

*b. 生成函数；三角余切和双曲余切的泰勒级数

伯努利数和伯努利多项式以一种漂亮的方式引导到余切和有关函数的泰勒展式. 这些展式最容易由所谓的伯努利函数的**生成函数**（generating function）得到，即由下面这个函数得到：

$$F(t, z) = \sum_{n=0}^{\infty} \phi_n(t) z^n. \tag{49}$$

这是一个 z 的幂级数，它的系数是参数 t 的伯努利函数. 根据方程(45)的傅立叶展开，对所有 t 和 $n \geqslant 2$ 我们有如下估计

$$|\phi_n(t)| \leqslant \left[\frac{2}{(2\pi)^n}\right] \sum_{k=1}^{\infty} \frac{1}{k^n} \leqslant \left[\frac{2}{(2\pi)^n}\right] \sum_{k=1}^{\infty} \frac{1}{k^2}$$

$$= \frac{\pi^2}{3(2\pi)^n} < \frac{4}{(2\pi)^n};$$

因此 $F(t, z)$ 的级数的 n 次项的绝对值小于 $4(|z|/2\pi)^n$. 如所周知，用级数

$$4\sum_{n=1}^{\infty}\left(\frac{|z|}{2\pi}\right)^n$$

作比较可得到：对所有的 t，z 的幂级数的收敛半径至少是 2π.

因为对 $|z| < 2\pi$ 内的一个固定的 z，$F(t, z)$ 的级数具有一个与 t 无关的收敛控制级数，由一般理论(见第 562 页)推知，对所有的 t，级数一致收敛. 所以它能够在这个区域内逐项积分；如果微分所得的级数也是一致收敛的话，它也能够逐项微分. 我们利用这个事实来确定 $F(t, z)$ 的明显的公式(见第 566 页). 在 $0 < t < 1$ 中，对 t 逐项微分(当 $t = 0$ 或 $t = 1$ 时，$\phi_1(t)$ 没有导数)就形式地得到

$$\frac{d}{dt}F(t, z) = \sum_{n=1}^{\infty} \psi_{n-1}(t)z^n$$

$$= z\sum_{n=1}^{\infty} \psi_{n-1}(t)z^{n-1}$$

$$= z\sum_{n=0}^{\infty} \psi_n(t)z^n$$

$$= zF(t, z).$$

这个级数同原来的级数具有相同形式，因而确实是一致收敛的，所以逐项微分是正当的. 因此，对于在 $|z| < 2\pi$ 内每一个固定的 z 和对于 $0 < t < 1$，生成函数 $F(t, z)$ 适合微分方程 $dF/dt = zF(t, z)$. 这个微分方程的通解是 $F = ce^{zt}$，其中 c 是一个因子，它的值依赖于参数 z（见第 237 页）. 为了确定 c，我们对 t 从 0 到 1 积分 $F(t, z)$ 的级数：

$$\int_0^1 F(t, z)dt = c\int_0^1 e^{zt}dt$$

$$= c\,\frac{e^z - 1}{z}$$

$$= \int_0^1 \sum_{n=0}^{\infty} z^n\psi_n(t)dt$$

$$= 1 + \sum_{n=1}^{\infty} z^n\int_0^1 \psi_n(t)dt = 1.$$

因而，$c = z/(e^z - 1)$，所以我们得到最后的结果

$$F(t, z) = \frac{ze^{zt}}{e^z - 1}. \tag{50}$$

在这个表达式中令 $t \to 0$，我们得到函数 $z/(e^z - 1)$ 的泰勒级数：

$$\lim_{t\to 0} F(t, z) = \frac{z}{e^z - 1} = 1 + \sum_{n=1}^{\infty} b_n z^n.$$

由于 $b_1 = -\frac{1}{2}$，两边加上 $\frac{1}{2}z$，就得到

$$\frac{z}{e^z-1}+\frac{z}{2}=1+\sum_{n=2}^{\infty}b_n z^n. \tag{51}$$

这个公式附带表明,数 $B_n^*=n!b_n$ 就是在第 590 页中介绍的伯努利数. 由于 $b_0=1$ 和当 n 为奇数时 $b_n=0$, 我们有

$$\begin{aligned}\frac{e^z+1}{e^z-1}\cdot\frac{z}{2}&=\frac{z}{2}\cdot\frac{e^{z/2}+e^{-z/2}}{e^{z/2}-e^{-z/2}}\\&=\frac{z}{2}\cdot\frac{2\cosh\frac{1}{2}z}{2\sinh\frac{1}{2}z}\\&=\frac{1}{2}z\coth\frac{1}{2}z=\sum_{n=0}^{\infty}b_{2n}z^{2n}\\&=\sum\frac{B_{2n}^*}{(2n)!}z^{2n}.\end{aligned} \tag{52}$$

因此,我们得到了已在第 591 页给出的双曲余切的泰勒级数;这些泰勒系数同伯努利数有简单的关系;现在我们已经证明了,对所有的 $|z|<2\pi$ 展式成立.

类似地, 我们得到通常的(三角的)余切的泰勒级数. 对 $|z|<2\pi$, $0<t<1$, 我们从对生成函数

$$G(t,z)=\sum_{n=0}^{\infty}(-1)^n\psi_{2n}(t)z^{2n}; \tag{53}$$

微分两次开始,我们得出 G 满足微分方程 $d^2G/dt^2+z^2G=0$. 这个方程的通解是 $G=a\cos(zt)+b\sin(zt)$, a 和 b 与 t 无关,但可能以 z 作为参数. 为了确定 a 和 b, 我们利用两个条件. 首先, 通过逐项积分求出 $\int_0^1 G(t,z)dt=1$; 其次, 对所有的 z, 通过逐项微分求出

$$\lim_{t\to0}\frac{dG(t,z)}{dt}=\frac{1}{2}z^2,$$

在微分中我们利用了当 $n>1$ 时

$$\psi'_{2n}(0)=\psi_{2n-1}(0)=b_{2n-1}=0.$$

这些条件蕴含着

$$a=\frac{z}{2}\cot\frac{z}{2},\qquad b=\frac{z}{2},$$

所以当 $|z|<2\pi$，$0<t<1$ 时有

$$G(t,z)=\frac{z}{2}\frac{\cos(zt-z/2)}{\sin(z/2)}.$$

详细的证明留给读者.

如果在这个公式中令 $t\to 0$，我们就得到在 $|z|<2\pi$ 内的余切的泰勒展式(见第 591 页)

$$G(0,z)=\sum_{n=0}^{\infty}(-1)^n b_{2n}z^{2n}=\frac{1}{2}z\cot\frac{1}{2}z. \tag{54}$$

c. 欧拉-马克劳林求和公式

在 5.4 b 节中,我们利用逐项分部积分得到了泰勒公式. 在下面欧拉的一个著名公式的类似推导中,伯努利多项式,或者确切地说,他们的周期开拓 $\psi_n(t)$ 代替了前面的多项式 $(t-b)^n/n!$.(我们用 0 和 1 代替了第 474 页上的 a 和 b，借助于把变量 t 变换为变量 s 的变换 $s=(t-a)/(b-a)$，这总是可能的,因而这是一个非本质的改变.)

我们以前的泰勒公式的推导是从关系式

$$f(1)-f(0)=\int_0^1 f'(t)dt \tag{55}$$

开始的,现在我们相应地改变为从关系式

$$\int_0^1 f(t)dt=\int_0^1 f(t)\psi_0(t)dt \tag{56}$$

开始. 这个公式引导出较大的对称性. 因为

$$\psi_0(t)=\psi_1'(t),\qquad \psi_1(+0)=-\frac{1}{2},$$

并且 $\psi_1(1-0)=\frac{1}{2}$，用分部积分公式

$$\int_0^1 u dv = uv\Big|_0^1 - \int_0^1 v du,$$

当 $u = f(t)$, $v = \psi_1(t)$, $f(0) = f_0$, $f(1) = f_1$, 得到(见第三章第 293 页)

$$\int_0^1 f(t)dt = \frac{1}{2}(f_0 + f_1) - \int_0^1 f'(t)\psi_1(t)dt,$$

或

$$\frac{1}{2}(f_0 + f_1) = \int_0^1 f(t)dt + \int_0^1 f'(t)\psi_1(t)dt, \tag{57}$$

这是关于左边的和与积分 $\int_0^1 f(t)dt$ 的偏差的一个显示表达式.

由于 $\psi_1(t)$ 的周期性, 相应的公式对每两个相继整数之间的区间保持有效,所以我们立即得到

$$\frac{1}{2}f_0 + f_1 + f_2 + \cdots + f_{n-1} + \frac{1}{2}f_n$$
$$= \int_0^n f(x)dx + \int_0^n f'(x)\psi_1(x)dx, \tag{58}$$

或对任何区间 $a \leqslant x \leqslant b$, a 和 b 为整数,我们有

$$f_a + f_{a+1} + \cdots + f_{b-1}$$
$$= \int_a^b f(x)dx + \int_a^b f'(x)\psi_1(x)dx - \frac{1}{2}(f_b - f_a). \tag{58a}$$

这样,我们就得到了关于左边的和(在增函数情况下就是内接矩形的面积)同右边的第一项(曲线下的面积)之间的差的一个精确的表达式;公式(58a)是欧拉-马克劳林求和公式最简明的表示.

很自然可反复用分部积分法去改善这个结果, 令 $u = f'(x)$, $dv = \psi_1(x)dx$, 将表达式 $\int_a^b f'(x)\psi_1(x)dx$ 积出来,我们得到

$$\int_a^b f'(x)\psi_1(x)dx = f'(x)\psi_2(x)\Big|_a^b - \int_a^b f''(x)\psi_2(x)dx.$$

由于

$$\psi_2(b) = \psi_2(a) = \psi_2(0) = b_2,$$

所以第一项形如

$$b_2[f'(b)-f'(a)];$$

第二项可再用分部积分法，得到

$$-b_3[f''(b)-f''(a)]+\int_a^b f'''(x)\phi_3(x)dx.$$

这里，由于 $b_3=0$，所以第一个表达式等于零；我们再用分部积分法，得到

$$b_4[f'''(b)-f'''(a)]-\int_a^b f''''(x)\phi_4(x)dx.$$

重复这种计算，一直进行到 ϕ_{2k}，我们就得到欧拉求和公式的一般形式

$$f_a+f_{a+1}+\cdots+f_{b-1}=\int_a^b f(x)dx-\frac{1}{2}[f(b)-f(a)]$$

$$+\sum_{n=1}^{k} b_{2n}[f^{(2n-1)}(b)-f^{(2n-1)}(a)]+R_k, \tag{59}$$

其中余项 R_k 可以写成下列两种形式中的任何一种：

$$R_k=-\int_a^b f^{(2k)}(x)\phi_{2k}(x)dx, \tag{60}$$

或

$$R_k=\int_a^b f^{(2k+1)}(x)\phi_{2k+1}(x)dx. \tag{60a}$$

d. 应用. 渐近表达式

收敛展开式. 欧拉的求和公式能够应用于不同的情况. 首先，如果当 $k\to\infty$ 时 $R_k\to 0$，则无穷级数

$$\sum_{n=1}^{\infty} b_{2n}[f^{(2n-1)}(b)-f^{(2n-1)}(a)]$$

收敛，因而公式给出了一个重要方法，把相应的级数的和表示成封闭形式，或者说把确定的函数表示成级数.

非收敛展开式. 其次，而且更重要的是，当 $k\to\infty$ 时余项 R_k 可以不趋于零；上述级数不一定收敛. 虽然最初可以出现当 k 增加

时绝对值 $|R_k|$ 减小，因而 $|R_k|$ 对适当选择的 k 值是很小的，然而 $|R_k|$ 在后来就开始（当 k 很大时）猛烈地增加. 在这种情况下，对数值计算来说，求和公式能够是一个重要的工具；虽然它不能像收敛级数那样获得任意高的精确度，但是我们仍能够在误差范围内计算左边的值，这个误差至多等于 $|R_k|$ 的最小值，而这最小值常常是一个令人很满意的精确度. 我们将考虑这两种现象都存在的例子.

例. 指数函数　我们首先对某个固定的 z 来考虑函数 $f(x)=e^{zx}$. 取 $a=0$ 和 $b=1$，对任何数 k，我们得到关系式

$$f_0=\int_0^1 f(x)dx-\frac{1}{2}[f(1)-f(0)]$$

$$+\sum_{n=1}^{k} b_{2n}[f^{(2n-1)}(b)-f^{(2n-1)}(a)]+R_k.$$

因此，

$$1=\frac{e^z-1}{z}-\frac{1}{2}(e^z-1)+\sum_{n=1}^{k} b_{2n}z^{2n-1}(e^z-1)+R_k$$

$$=\frac{e^z-1}{z}\cdot\left[1-\frac{z}{2}+\sum_{n=1}^{k} b_{2n}z^{2n}\right]+R_k,$$

式中

$$R_k=-\int_0^1 z^{2k}e^{zx}\psi_{2k}(x)dx.$$

由于 $|\psi_{2k}(x)|\leqslant 4/(2\pi)^{2k}$ （第 657 页），从而得到

$$|R_k|\leqslant |z|^{2k}\cdot e^{|z|}\frac{4}{(2\pi)^{2k}}=4e^{|z|}\left(\frac{|z|}{2\pi}\right)^{2k},$$

或者至少对 $|z|<2\pi$ 而言 $R_k\to 0$. 因此，在求和公式中，对于这些 z 的值我们可以容许 k 无限地增长，就对函数 $z/(e^z-1)$ 得到

$$\frac{z}{e^z-1}=1-\frac{1}{2}z+\sum_{n=1}^{\infty} b_{2n}z^{2n}, \tag{61}$$

这个公式已用另外方法求出过（第 659 页）. 我们注意，收敛区间

仍是 $|z|<2\pi$.

e. 幂级数的和;伯努利数的递推公式

当收敛的欧拉求和公式右边的级数仅含有限多项，特别是如果函数 $f(x)$ 是一个 $r\geqslant 1$ 阶的多项式,从而 $f^{r+1}(x)$ 恒等于0时,就呈现出更为简单的情况.我们选择 $f(x)=x^r$, $a=0$, $b=n$, 并且 $k>\frac{1}{2}r$. 为简单起见，我们再次引入以前（第 659 页）定义的伯努利数的序列 B_n^*, 对所有的 n 定义为 $B_n^*=n!b_n$. 注意到

$$B_0^*=1,\ B_1^*=-\frac{1}{2},\ B_3^*=B_5^*=B_7^*=\cdots=B_{2n+1}^*=0,$$

我们看出,欧拉公式 (59) 取下述形式

$$\begin{aligned}
&1+2^r+3^r+\cdots+(n-1)^r\\
&=\int_0^n x^r dx+\sum_{\nu=1}^{r}\frac{B_\nu^*}{\nu!}\cdot(f^{(\nu-1)}(n)-f^{(\nu-1)}(0))\\
&=\frac{n^{r+1}}{r+1}+\sum_{\nu=1}^{r}\frac{B_\nu^*}{\nu!}r(r-1)\cdots(r-\nu+2)n^{r-\nu+1}\\
&=\frac{1}{r+1}\left\{n^{r+1}+\sum_{\nu=1}^{r}\binom{r+1}{\nu}B_\nu^* n^{(r+1)-\nu}\right\}\\
&=\frac{1}{r+1}\left\{\sum_{\nu=0}^{r+1}\binom{r+1}{\nu}n^{(r+1)-\nu}B_\nu^*-B_{r+1}^*\right\}.
\end{aligned}$$

这个公式能够用符号写成下述形式

$$\begin{aligned}
&1+2^r+3^r+\cdots+(n-1)^r\\
&=\frac{1}{r+1}\{(n+B^*)^{r+1}-B^{*r+1}\},
\end{aligned}\tag{62}$$

式中括号里的项要在形式上用二项式定理展开，并且诸“幂” B^{*k} 的每一个要用相应的伯努利数 B_k^* 代替. 例如

$$\begin{aligned}
&1+2^2+3^2+\cdots+(n-1)^2\\
&=\frac{1}{3}(n^3+3n^2B_1+3nB_2)
\end{aligned}$$

$$= \frac{1}{6}(2n^3 - 3n^2 + n),$$

$$1 + 2^3 + 3^3 + \cdots + (n-1)^3$$

$$= \frac{1}{4}n^2(n-1)^2$$

(参看第 60 页)。

令 $n = 1$，公式 (62) 就取下述形式

$$\frac{1}{r+1}\{(1+B^*)^{r+1} - B^{*r+1}\} = 0,$$

或

$$(1+B^*)^{r+1} = B^{*r+1}, \quad r \geqslant 1. \tag{62a}$$

这正好是在第 590 页给出的 B_k^* 的递推公式。

f. 欧拉常数和斯特林级数

在第二种情况下，即在发散的情况下，欧拉-马克劳林公式应用的一个例是由函数 $f(x) = \frac{1}{x}$ 给出的。在公式中取 $a = 1$，$b=n$，则根据 (58a) 有

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n-1}$$

$$= \int_1^n \frac{dx}{x} + \frac{1}{2}\left(1 - \frac{1}{n}\right) - \int_1^n \frac{\psi_1(x)}{x^2}\,dx$$

$$= \log n + \frac{1}{2} - \frac{1}{2n} - \int_1^n \frac{\psi_1(x)}{x^2}\,dx, \tag{63}$$

或

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{n} - \log n$$

$$= \frac{1}{2} + \frac{1}{2n} - \int_1^n \frac{\psi_1(x)}{x^2}\,dx.$$

因为对所有的 x 都有 $|\psi_1(x)| \leqslant \frac{1}{2}$，所以当 $n \to \infty$ 时，右

边的积分收敛;于是这个积分的绝对值总比收敛积分 $\int_1^\infty dx/x^2$ 的值小. 因此,我们得到下述关系式

$$\lim_{n\to\infty}\left[\sum_{k=1}^{n}\frac{1}{k}-\log n\right]=\frac{1}{2}-\int_1^\infty\frac{\psi_1(x)}{x^2}\,dx=C, \qquad (64)$$

确定的常数 C,即欧拉常数,已在第 553 页中介绍过.

于是我们有两个结果: 调和级数增长的阶和对数增长的阶相同,两者都发散到无穷,而二者之间的差有一个明显的表达式

$$\sum_{k=1}^{\infty}\frac{1}{k}-\log n-C=R_n=\frac{1}{2n}+\int_n^\infty\frac{\psi_1(x)}{x^2}\,dx.$$

我们注意到,当 $n\to\infty$ 时, R_n 至少是一阶地趋于 0.

当我们在第 662 页的公式 (59) 中令 $f(x)=\log x$, $a=1$, $b=n$, 就得到一个更重要的应用. 这时

$$\begin{aligned}
&\log 1+\log 2+\cdots+\log(n-1)\\
&=n\log n-n+1-\frac{1}{2}\log n\\
&\quad-\sum_{m=1}^{k}b_{2m}(2m-2)!\left(1-\frac{1}{n^{2m-1}}\right)\\
&\quad+\int_1^n\frac{(2k)!}{x^{2k+1}}\psi_{2k+1}(x)dx.
\end{aligned}$$

两边加上 $\log n$, 就得到

$$\log n!=\left(n+\frac{1}{2}\right)\log n-n+c_k+\sum_{m=1}^{k}\frac{(2m-2)!}{n^{2m-1}}b_{2m}-r_k(n), \qquad (65)$$

式中

$$c_k=1-\sum_{m=1}^{k}b_{2m}(2m-2)!+\int_1^\infty\frac{(2k)!}{x^{2k+1}}\psi_{2k+1}(x)dx,$$

$$r_k(n)=\int_n^\infty\frac{(2k)!}{x^{2k+1}}\psi_{2k+1}(x)dx.$$

由于函数 $\psi_{2k+1}(x)$ 是周期的，因而对所有的 x 是有界的，所以当 $k>0$ 时，上面的广义积分收敛（见第 326 页）. 如果我们根据(65)，注意到当 $n\to\infty$ 时

$$c_k=\lim_{n\to\infty}\log\left(\frac{n!e^n}{n^{n+1/2}}\right),$$

我们就能够求出常数 c_k 的值. 于是我们从第 531 页的斯特林公式（14）（或如在第 298 页那样直接从关于 π 的瓦里斯乘积推出）推出 $c_k=\log\sqrt{2\pi}$. 如果我们仍把伯努利数 b_{2m} 表示成 $(-1)^{m-1}B_m/(2m)!$（见第 656 页公式（47）），我们就得到所谓**斯特林级数**

$$\log\left(\frac{n!}{\sqrt{2\pi}\,n^{n+1/2}e^{-n}}\right)=\sum_{m=1}^{k}\frac{(-1)^{m-1}B_m}{2m(2m-1)n^{2m-1}}-r_k(n).$$

这个公式是斯特林公式的一种精炼化. 对任何一个固定的正整数 k 和大的 n，这个和中的项趋于 0，其阶分别为 $1/n, 1/n^3, 1/n^5, \cdots, 1/n^{2k-1}$. 因为 $\psi_{2k+1}(x)$ 是一个有界函数，所以余项 $r_k(n)$ 像 $1/n^{2k}$ 一样接近于零. 于是对于固定的 k 和很大的 n，这个和中的每一项比后面项大得多，因而余项比和中的所有的项都要小. 因此，我们得到如下形式的近似公式

$$\begin{aligned}\log\left(\frac{n!}{\sqrt{2\pi}n^{n+1/2}e^{-n}}\right)&=\frac{B_1}{1\cdot2}\frac{1}{n}-\frac{B_2}{3\cdot4}\frac{1}{n^3}+\frac{B_3}{5\cdot6}\frac{1}{n^5}-+\cdots\\&=\frac{1}{12}\frac{1}{n}-\frac{1}{360}\frac{1}{n^3}+\frac{1}{1260}\frac{1}{n^5}\\&\quad-\frac{1}{1680}\frac{1}{n^7}+\frac{1}{1188}\frac{1}{n^9}-+\cdots.\end{aligned}\tag{66}$$

无论如何在同样的见解下，这个展开式绝不能作为收敛的无穷级数来看待. 它只是在下面的意义上**渐近地**（asymptotically）正确，即如果我们在某一固定的项数之后，比如说 k 项之后断开这级数，则当 n 足够大时，误差 r_k 同所有剩下的项相比是小的. 对固

定的 n，我们决不能够通过取越来越多的项来使误差任意地小。实际上，无穷级数(66)是发散的，这是我们从第 656 页对伯努利数的估计可以直接看出来的. 对于一个给定的大的 n，这级数有一个人们可能利用的最优的项数. 例如，对于适当大的数 n，就有近似式

$$n! \approx \sqrt{2\pi} n^{n+1/2} e^{n+1/12n};$$

对于很大的 n，公式

$$n! \approx \sqrt{2\pi} n^{n+1/2} e^{n+1/12n-1/360n^3}$$

给出了一个更精确的近似式，等等.

问　　题

8.1 节，第 601 页

1. 周期函数 f 的基本周期 T 定义为 f 的正周期的最大下界. 证明:

(a) 如果 $T \neq 0$，则 T 是一个周期;

(b) 如果 $T \neq 0$，则每一个别的周期是 T 的整数倍;

(c) 如果 $T = 0$ 并且 f 在每一点都是连续的，则 f 是一常数函数.

2. 证明: 如果 f 有不可通约的周期 T_1 和 T_2，则基本周期为零. 给出一个非常数函数拥有不可通约周期的例.

3. 令 f 和 g 分别有基本周期 a 和 b. 如果 a 和 b 是可通约的，也就是说 $a/b = q/p$，其中 p 和 q 是互素的整数，则用一例表明，$f + g$ 可有任何值 m/n 作为它的基本周期，其中 $m = aq = bp$ 并且 n 是任何自然数.

8.5 节，第 629 页

1. 对于在区间 $0 \leqslant x \leqslant 1$ 上的函数 $f(x) = \pi x$，求它的一个纯正弦级数和一个纯余弦级数的傅立叶级数.

2. 说明如何把一个定义在任意有界区间上的函数表示为傅立叶级数。

3. 根据关系式 $\cos \pi x = \dfrac{\sin 2\pi x}{2 \sin \pi x}$ 求余弦的无穷乘积.

4. 利用正弦和余弦的无穷乘积，计算

(a) $\dfrac{2}{1} \cdot \dfrac{2}{3} \cdot \dfrac{6}{5} \cdot \dfrac{6}{7} \cdot \dfrac{10}{9} \cdot \dfrac{10}{11} \cdot \dfrac{14}{13} \cdots;$

(b) $2 \cdot \dfrac{2}{3} \cdot \dfrac{4}{3} \cdot \dfrac{8}{9} \cdot \dfrac{10}{9} \cdot \dfrac{14}{15} \cdot \dfrac{16}{15}.$

5．把双曲余切．表示成部分分式

6．确定 $f(x)=f(\pi-x)$ 的偶函数和奇函数的傅立叶展开式的系数的特殊性质.

8.6 节，第 636 页

1．研究函数 $-\log 2\left|\sin\frac{x}{2}\right|$ 的傅立叶展开式

$$\cos x+\frac{\cos 2x}{2}+\frac{\cos 3x}{3}+\cdots$$

的收敛性.

8.7 节，第 641 页

1．对可有若干间断点的分段光滑函数 f 证明巴色瓦等式.

附录 II.1，第 654 页

1．证明

$$\phi_n(t)=\frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}B_k^*t^{n-k}.$$

2．对 $n\geqslant 1$ 证明

$$\phi_n(t)=(-1)^n\phi_n(1-t).$$

3．利用余切的部分分式表达式，把 $\pi x\cot\pi x$ 展为 x 的幂级数. 同第661页给的级数进行比较，证明

$$\sum_{\nu=1}^{\infty}\frac{1}{\nu^{2m}}=(-1)^{m-1}\frac{(2\pi)^{2m}}{2\cdot(2m)!}B_{2m}^*.$$

4．证明

$$\sum_{\nu=1}^{\infty}\frac{1}{(2\nu-1)^{2m}}=\frac{(-1)^{m-1}(2^{2m}-1)\pi^{2m}}{2(2m)!}B_{2m}^*.$$

5．证明

$$\sum_{\nu=1}^{\infty}\frac{(-1)^{\nu}}{\nu^{2m}}=\frac{(-1)^m(2^{2m}-2)\pi^{2m}}{2\cdot(2m)!}B_{2m}^*.$$

6．利用正弦和余弦的无穷乘积，证明

(a) $\log\left(\frac{\sin x}{x}\right)=-\sum_{\nu=1}^{\infty}\frac{(-1)^{\nu-1}2^{2\nu-1}B_{2\nu}^*}{(2\nu)!\,\nu}x^{2\nu}$;

(b) $\log\cos x = -\sum_{\nu=1}^{\infty}\frac{(-1)^{\nu-1}2^{2\nu-1}(2^{2\nu}-1)B_{2\nu}^{*}}{(2\nu)!\,\nu}x^{2\nu}$

7. 证明

(a) $\int_0^1 \frac{\log x}{1-x}\,dx = -\frac{\pi^2}{6}$;

(b) $\int_0^1 \frac{\log x}{1+x}\,dx = -\frac{\pi^2}{12}$.

第九章　关于振动的最简单类型的微分方程

前面曾有几处我们遇到过**微分方程**（differential equation），也就是这样的方程，有一个未知函数将由之确定，并且这个方程不仅仅包含这个函数本身，而且也包含它的导数.

这种类型的最简单问题是求一个给定的函数 $f(x)$ 的不定积分：求满足微分方程 $y'-f(x)=0$ 的函数 $y=F(x)$. 此外，在第三章里，第 237 页，我们已经说明过，形如 $y'=\alpha y$ 的方程为指数函数 $y=ce^{\alpha x}$ 所满足，我们还用微分方程(第 334 页)刻划了三角函数. 正如在第四章里(例如第 429 页)所看到的那样，微分方程的起源与力学问题有关，并且纯数学的很多分支和大多数的应用数学都的确依赖于微分方程. 在这章里，不去涉及一般的理论，我们将考虑关于振动的最简单类型的微分方程. 这些不仅仅有理论价值而且在应用数学中也是非常重要的.

记住下面一般概念和定义将是方便的. 微分方程的**解**是指这样一个函数，当我们把它代入微分方程后，对一切所考虑的自变量值来说它"同样"满足这方程. 经常用**积分**这个术语代替**解**；首先，因为这个问题或多或少是一般的积分问题的推广；其次，因为常常碰到的是，解实际上是由积分求得的.

9.1　力学和物理学的振动问题

a. 简单的机械振动

机械振动的最简单的类型已经在第四章(第 428 页)研究过. 在那里我们研究了一个质量为 m 的质点，它在 x 轴上自由运动，并且被一个恢复力把它拉回到它的初始位置 $x=0$. 我们把这个恢复力的大小取为与位移 x 成比例，实际上，它等于 $-kx$，其中 k 是一个正的常数，而负号表示这个力总是指向原点. 我们现在假定还有摩

擦力，这个摩擦力与质点的速度 $dx/dt = \dot{x}$ 成比例，但与之反向．于是这个力由形如 $-r\dot{x}$ 的表达式给出，式中带有一个正的摩擦常数 r．最后，我们将假定质点受到外力的作用，外力是时间 t 的函数 $f(t)$．于是根据牛顿基本定律，质量 m 和加速度 $\ddot{x}$ 的乘积必定等于总力，即等于弹性力加摩擦力再加外力．这可以用方程

$$m\ddot{x} + r\dot{x} + kx = f(x) \tag{1}$$

表示．

这个方程支配着质点的运动．如果我们回想起前面的微分方程的例题，例如 $\dot{x} = dx/dt = f(t)$ 的积分问题用 $x = \int f(t)dt + c$ 求解，或用第 428 页上特殊微分方程 $m\ddot{x}+kx=0$的解法，我们注意到，这些问题具有无穷个相异的解．这里我们也将发现有无穷个解，这个事实是用下面的方法表示出来的．寻求这个微分方程的一般解或完全积分 $x(t)$ 是可能的，它的一般解不仅依赖于自变量 t 而且也依赖于两个任意参数 c_1 和 c_2，c_1 和 c_2 称为积分常数．把这些常数指定特殊值我们就得到一个特解，并且每一个解都能够通过对这些常数指定特殊的值而得到．

这个事实是完全可以理解的（也参看第 428 页）．我们不能期望单独由微分方程就能完全确定这个运动．然而，我们可以合理地设想，在一个给定瞬间，譬如说在时刻 $t=0$，我们应该能够任意选择初始位移 $x(0)=x_0$ 和初始速度 $\dot{x}(0)=\dot{x}_0$（简称为初始状态）；换言之，我们应该能够在时刻 $t=0$ 给质点以任何初始位置和初始速度使之运动．这样做了之后，我们可以期待这个运动的其余一切都完全被确定了．在一般解里的两个任意常数 c_1 和 c_2 刚好足够使我们能够选择适合这些初始条件（initial condition）的特解．在下节中我们将看到仅能以一种方式这样做．

如果没有出现外力，即如果 $f(t)=0$，则称运动为自由运动，这个微分方程就说是齐次的．如果 $f(t)$ 不是对一切 t 值来说都等于零，我们就说运动是强迫的，并且称微分方程是非齐次的．偶尔也把 $f(t)$ 项叫做扰动项．

b．电的振荡

上面所描述的这一简单的机械系统在物理上只能近似地实现．摆就是一个例，只要它的摆动是小的．一个磁针的振动，电话或

扩音器膜片中心的振动，以及其它机械振动都是在一定精确程度内能够用如我们所描述的那样的系统来表示．但有另外一类现象在很大的精确度上相当于我们的微分方程(1).这就是振荡电路．

我们来考察画在图9.1中的电路简图,在这个电路中有电感μ,电阻ρ和电容$C=1/\kappa$.我们还假定有一个外电动势$\phi(t)$作用在线路上,它是时间t的已知函数，例如由发电机供给的电压或由电波得到的电压.为了描述在线路里发生的过程，我们用E表示电容器的电压，用Q表示电容器内的电荷量．于是这些量可以用等式$CE=E/\kappa=Q$联系起来．电流I与电压E一样是时间的函数，它定义为每单位时间内电荷量的变化率，亦即，它作为电容器上电量减少的变化率：$I=-\dot{Q}=-dQ/dt=-\dot{E}/\kappa$．欧姆定律指出电流和电阻的乘积等于电动势（电压)；亦即，它等于电容器的电压E减去由于自感产生的反电动势，再加上外电动势$\phi(t)$．于是我们得到方程$I\rho=E-\mu\dot{I}+\phi(t)$或$-(\rho/\kappa)\dot{E}=E+(\mu/\kappa)\ddot{E}+\phi(t)$，亦即，线路里的电压满足方程$\mu\ddot{E}+\rho\dot{E}+\kappa E=-\kappa\phi(t)$．因此我们得到了恰好是类型(1)的微分方程．代替质量我们有电感，代替摩擦力有电阻，并且代替弹性常数的有电容的倒数，而外电动势（不计常数因子）相当于外力．如果电动势为零，则微分方程是齐次的．

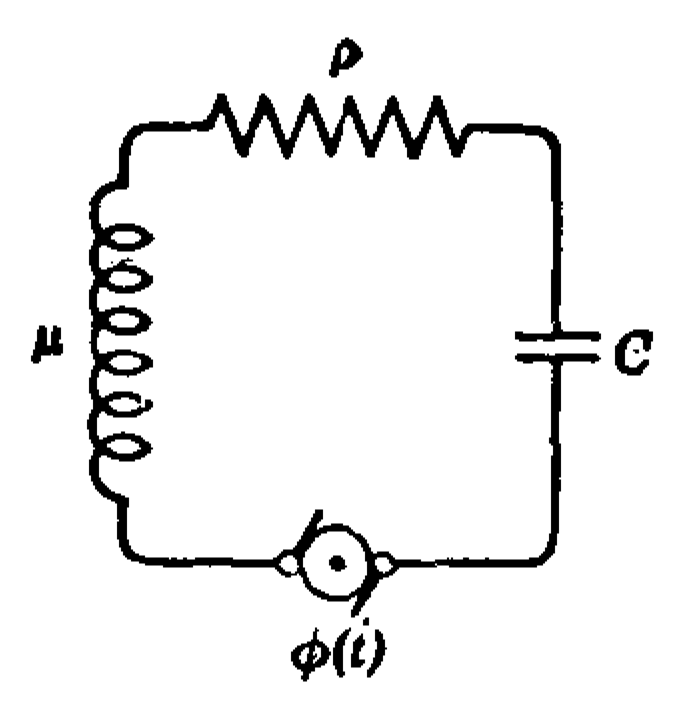


图9.1　振荡电路

如果我们用$-1/\kappa$乘微分方程的两边并且对于时间进行微分，我们就得到电流I所满足的方程

$$\mu\ddot{I}+\rho\dot{I}+\kappa I=\dot{\phi}(t),$$

它与电压方程比较仅右边不同，而对于自由振动$(\phi=0)$来说则具有完全相同的形式．

9.2　齐次方程的解法．自由振动

a．形式解

我们能够容易得到齐次方程(1) $m\ddot{x}+r\dot{x}+kx=0$的一个

形式为指数表达式的解，通过决定常数 λ 使表达式 $e^{\lambda t}=x$ 是一个解．如果我们在微分方程里代入这个试验性质的解和它的导数 $\dot{x}=\lambda e^{\lambda t}$，$\ddot{x}=\lambda^2 e^{\lambda t}$，并消去公因子 $e^{\lambda t}$，我们就得到关于 λ 的二次方程

$$m\lambda^2+r\lambda+k=0. \tag{2}$$

这个方程的根是

$$\lambda_1=-\frac{r}{2m}+\frac{1}{2m}\sqrt{r^2-4mk},$$

$$\lambda_2=-\frac{r}{2m}-\frac{1}{2m}\sqrt{r^2-4mk}.$$

两个式子 $x=e^{\lambda_1 t}$ 和 $x=e^{\lambda_2 t}$ 中的每一个，至少在形式上是微分方程的一个特殊解，这只要我们反回去算一算就会明白的．现在可能发生三种不同情况：

1. $r^2-4mk>0$．这时，两根 λ_1 和 λ_2 是实的，负的，和不相等的，从而我们得到微分方程的两个解，

$$u_1=e^{\lambda_1 t} \quad 和 \quad u_2=e^{\lambda_2 t}.$$

借助于这两个解，我们能够立即构造一个含有两个任意常数的解．因为经过微分之后我们看到

$$x=c_1u_1+c_2u_2 \tag{3}$$

也是微分方程的一个解．在 9.3 节里我们将证明这个式子事实上是方程的最一般的解；亦即，只要把 c_1 和 c_2 代入适当数值就能得到方程的**每一个**解．

2. $r^2-4mk=0$．这时二次方程有重根．于是，若暂且不计常数因子，我们在开始仅得到一个解 $x=w_1=e^{-rt/2m}$．但是在这种情形下，我们容易验证函数

$$x=w_2=te^{-rt/2m}$$

也是微分方程的一个解[1)]．因为，我们求得

1) 我们通过下述的极限过程自然地导出这个解：若 $\lambda_1 \neq \lambda_2$ 则表达式 $(e^{\lambda_1 t}-e^{\lambda_2 t})/(\lambda_1-\lambda_2)$ 也是一个解．我们现在令 λ_1 趋于 λ_2 并用 λ 代替 λ_1，λ_2，这个式子就变成 $d(e^{\lambda t})/d\lambda=te^{\lambda t}$．

$$\dot{x} = \left(1 - \frac{r}{2m}t\right)e^{-rt/2m}, \quad \ddot{x} = \left(\frac{r^2}{4m^2}t - \frac{r}{m}\right)e^{-rt/2m},$$

直接代入就看到它满足微分方程

$$m\ddot{x} + r\dot{x} + \frac{r^2}{4m}x = m\ddot{x} + r\dot{x} + kx = 0.$$

于是式子

$$x = c_1 e^{-rt/2m} + c_2 t e^{-rt/2m} \tag{4}$$

重新给出微分方程的一个含有两个任意积分常数 c_1 和 c_2 的解.

3. $r^2 - 4mk < 0$. 我们设 $r^2 - 4mk = -4m^2\nu^2$ 就得到微分方程的两个复形式的解，由表达式 $x = u_1 = e^{-rt/2m+i\nu t}$ 和 $x = u_2 = e^{-\nu t/2m-i\nu t}$ 给出. 用欧拉公式

$$e^{\pm i\nu t} = \cos\nu t \pm i\sin\nu t$$

得到复解 u_1 的实部和虚部，一方面有表达式

$$v_1 = e^{-rt/2m}\cos\nu t, \quad v_2 = e^{-rt/2m}\sin\nu t;$$

另方面又有表达式

$$v_1 = \frac{u_1 + u_2}{2}, \qquad v_2 = \frac{u_1 - u_2}{2}.$$

从这第二种表达形式我们看到 v_1 和 v_2 都是微分方程的（实）解. 这可以直接通过求微分和代入来证实，我们把它做为一个简单习题留给读者.

从这两个特解我们又能构成一个一般解

$$x = c_1 v_1 + c_2 v_2 = (c_1\cos\nu t + c_2\sin\nu t)e^{-rt/2m}, \tag{5}$$

其中含有两个任意常数 c_1 和 c_2. 也可以把它写成这样的形式

$$x = ae^{-rt/2m}\cos\nu(t - \delta), \tag{6}$$

在这里，我们写成了 $c_1 = a\cos\nu\delta$, $c_2 = a\sin\nu\delta$, 而 a, δ 是两个新的常数.

我们记起了我们曾经遇到过 $r = 0$ 时的这个特殊情形（5.4节）的解.

b. 解的诠释

在 $r > 2\sqrt{mk}$ 和 $r = 2\sqrt{mk}$ 两种情形下，解是用指数曲线或用函数 $te^{-rt/2m}$ 的图象给出的，或者是用这些曲线的叠加给出的，而函数 $te^{-rt/2m}$ 对于大的 t 值来说与指数曲线相类似. 在这些情形下过程是非周期的；亦即，当时间增长时，"距离" x 逐渐地接近于值 0，在 $x=0$ 附近无振动. 因此运动没有振动. 摩擦或阻尼的作用是如此之大以至它阻止了弹性力产生振动运动.

当 $r < \sqrt{2mk}$ 时却完全不同，这里阻尼很小以至产生了复根 λ_1，λ_2. 公式 $x = a\cos\nu(t-\delta)e^{-rt/2m}$ 给出受阻尼的简谐振动. 这些是服从正弦法则和有圆频率 $\nu = \sqrt{k/m - r^2/4m^2}$ 的振动，但是它的振幅是由 $ae^{-rt/2m}$ 给出的，不再是常数了. 这就是说，振幅按指数减小；$r/2m$ 越大，缩小的速率越快. 在物理文献中，常常把这个阻尼因子称为这个阻尼振动的衰减常数，这个术语指出了振幅的对数的减小率为 $r/2m$. 图 9.2 说明了这种类型的阻尼振动. 如同前面一样，我们称量 $T=2\pi/\nu$ 是振动的周期而量 $\nu\delta$ 是位相的位移. 对于特殊情形 $r=0$，我们再一次得到具有频率为 $\nu_0 = \sqrt{k/m}$ 的简谐振动，这个频率是无阻尼振动系统的自然频率.

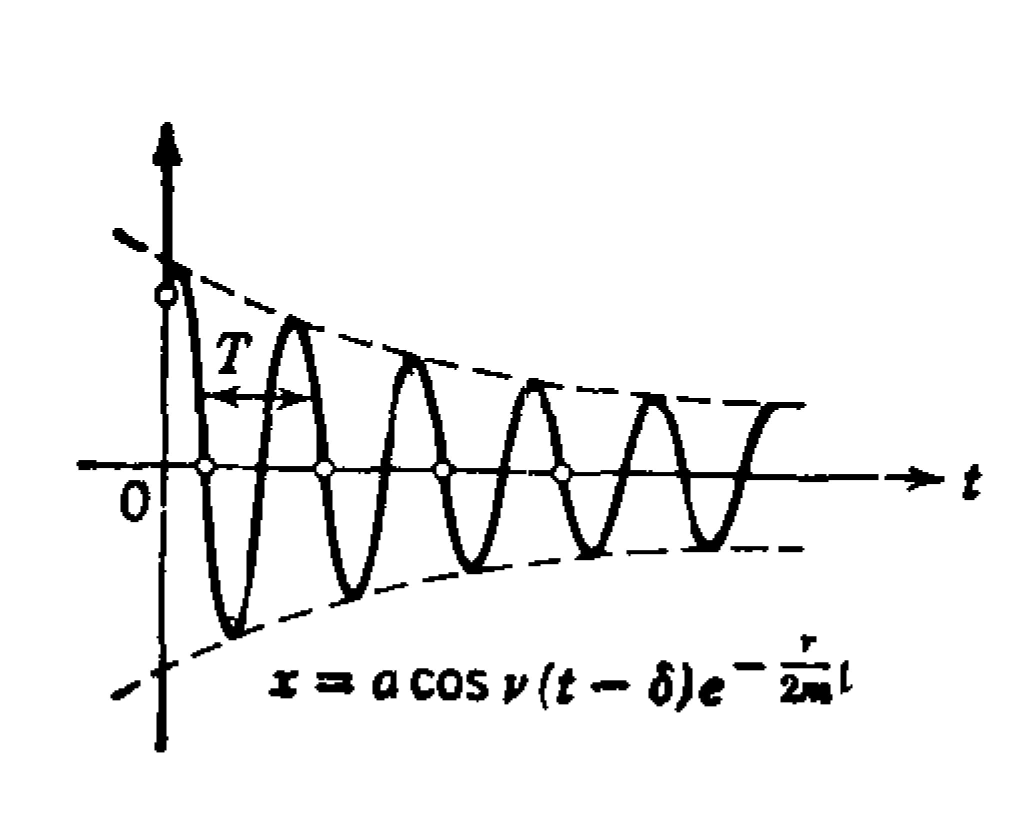


图 9.2 阻尼简谐振动

c. 满足给定的初始条件. 解的唯一性

我们还需要证明具有两个常数 c_1 和 c_2 的解能够适合于任意预先指定的初始状态，并且它代表了方程的所有可能的解. 假定我们要寻找一个解，它在时间 $t=0$ 满足初始条件 $x(0)=x_0$，$\dot{x}(0)=\dot{x}_0$，这里数 x_0 和 $\dot{x}_0$ 可以是任意值. 则在 9.2a 节（第 674 页）的第 1 种情形下，我们需令

$$c_1 + c_2 = x_0,$$
$$c_1\lambda_1 + c_2\lambda_2 = \dot{x}_0.$$
这样,对于常数 c_1 和 c_2 相应地有两个线性方程,并且它们有唯一的解

$$c_1 = \frac{\dot{x}_0 - \lambda_2 x_0}{\lambda_1 - \lambda_2}, \qquad c_2 = \frac{\dot{x}_0 - \lambda_1 x_0}{\lambda_2 - \lambda_1}.$$

在第 2 种情形下,用同样手续得到两个线性方程

$$c_1 = x_0,$$
$$\lambda c_1 + c_2 = \dot{x}_0 \qquad \left(\lambda = -\frac{r}{2m}\right),$$

从这里又能唯一地决定 c_1 和 c_2. 最后,在第 3 种情形下,决定这些常数的方程所取的形式为

$$a\cos\nu\delta = x_0,$$
$$a\left(\nu\sin\nu\delta - \frac{r}{2m}\cos\nu\delta\right) = \dot{x}_0,$$

具有解

$$\delta = \frac{1}{\nu}\arccos\frac{x_0}{a}, \quad a = \frac{1}{\nu}\sqrt{\left[\nu^2 x_0^2 + \left(\dot{x}_0 + \frac{r}{2m}x_0\right)^2\right]}.$$

这样一来,我们就证明了一般解能够使它适合任何的初始条件. 我们还需要证明,再也没有其它的解. 为此我们只需要证明,对于一个给定的初始状态,决不可能有两个不同的解.

如果存在两个这样的解 $u(t)$ 和 $v(t)$, 使得 $u(0) = x_0$, $\dot{u}(0) = \dot{x}_0$ 并且 $v(0) = x_0$, $\dot{v}(0) = \dot{x}_0$, 则它们的差 $w = u - v$ 也是微分方程的一个解,并且我们有 $w(0) = 0$, $\dot{w}(0) = 0$. 这个解是对应于静止的初始状态,亦即对应的状态是,当时间 $t = 0$ 时质点在静止的位置上并且有零速度. 我们必须证明它决不会使它自己运动. 为此我们用 $2\dot{w}$ 乘微分方程 $m\ddot{w} + r\dot{w} + kw = 0$ 的两边并且记住 $2\dot{w}\ddot{w} = (d/dt)\dot{w}^2$ 和 $2w\dot{w} = (d/dt)w^2$. 于是我们得到

$$\frac{d}{dt}(m\dot{w}^2) + \frac{d}{dt}(kw^2) + 2r\dot{w}^2 = 0.$$

如果我们在时刻 $t=0$ 和 $t=\tau$ 之间积分，并且使用初始条件 $w(0)=0$，$\dot{w}(0)$，我们就有

$$m\dot{w}^2(\tau)+kw^2(\tau)+2r\int_0^\tau\left(\frac{dw}{dt}\right)^2dt=0.$$

然而，假若在任何一个时刻 $\tau>0$ 时函数 w 不为0，则这个方程将会导出矛盾. 因为我们已经取 m，k 和 r 为正值，所以方程左边为正，然而右边为零. 因此 $w=u-v$ 总是等于 0，这就证明了我们的解是唯一的.

9.3 非齐次方程. 强迫振动

a. 一般说明. 叠加法

在求解出现外力 $f(t)$ 的问题之前，也就是在求解非齐次方程之前，我们作如下的说明.

若 w 和 v 是非齐次方程的两个解，则差 $u=w-v$ 满足齐次方程；只要代入即可看出. 反之，若 u 是齐次方程的解而 v 是非齐次方程的解，则 $w=u+v$ 也是非齐次方程的解. 因此从非齐次方程的一个解[1]加上齐次方程的完全积分，我们得到非齐次方程的全部解. 因此我们只需要找到非齐次方程的一个解就可以了. 这在物理上意味着：如果我们有一个由于外力引起的强迫振动，在它之上叠加一个用齐次方程的一个解所表示的一个任意的自由振动，那末我们就得到一个现象，它如同原来的强迫振动一样满足相同的非齐次方程. 若出现摩擦力，则在振动运动情形下的自由振动由于阻尼因子 $e^{-rt/2m}$ 的缘故，必定随时间的继续而衰减. 因此对于给定的带有摩擦的强迫振动，我们叠加什么样的自由振动都是无关紧要的；当时间继续下去时，运动总是趋向于相同的最终状态.

其次，我们注意到，可以把力 $f(t)$ 进行等效的分解. 这意思就是指：若 $f_1(t)$，$f_2(t)$ 和 $f(t)$ 是满足

$$f_1(t)+f_2(t)=f(t)$$

1) 常称为特殊积分或特解.

的三个函数，并且若 $x_1 = x_1(t)$ 是微分方程 $m\ddot{x} + r\dot{x} + kx = f_1(t)$ 的解，而 $x_2 = x_2(t)$ 是微分方程 $m\ddot{x} + r\dot{x} + kx = f_2(t)$ 的解，则 $x(t) = x_1(t) + x_2(t)$ 是微分方程

$$m\ddot{x} + r\dot{x} + kx = f(t) \tag{7}$$

的解。当 $f(t)$ 包括任何数目的项时，相应的叙述当然保持成立。这个简单但是重要的事实称为**叠加原理**。它的证明可以从方程一眼看出。把函数 $f(t)$ 分成两项或更多的项，我们就能够把微分方程分解成几个方程，它们在有些情况下可以比较容易处理。

最重要的是具有周期外力 $f(t)$ 的情况。这样的周期外力可以借助傅立叶级数展式分解成纯周期的分量，因此[1]能够用有限个纯周期函数的和去逼近它，并能达到我们所希望的接近程度。因此只需要求解右边具有形状为

$$a\cos\omega t \quad 或 \quad b\sin\omega t,$$

的微分方程就可以了，这里 a, b 和 ω 是任意常数。

若使用复的符号来代替这些三角函数，则能得到更为简洁的解。我们设 $f(t) = ce^{i\omega t}$，叠加原理说明我们只需考虑微分方程

$$m\ddot{x} + r\dot{x} + kx = ce^{i\omega t}, \tag{8}$$

其中 c 是任意的实或复的常数。这样的一个微分方程实际上代表了两个实的微分方程。我们如果把右边分裂成两项，例如取 $c=1$ 并且写成 $e^{i\omega t} = \cos\omega t + i\sin\omega t$，则由两个实微分方程 $m\ddot{x} + r\dot{x} + kx = \cos\omega t$ 和 $m\ddot{x} + r\dot{x} + kx = \sin\omega t$ 的解 x_1, x_2 的组合而构成了这个复微分方程的解 $x = x_1 + ix_2$。反之，如果我们首先求出复形式的微分方程的解，那么这个解的实部给出函数 x_1，虚部给出函数 x_2。

b. 非齐次方程的解法

我们用直观自然地启发的一种方案来求解微分方程(8)。设 c 是实的，并且(暂时假定) $r \neq 0$。现在我们想像；存在一个运动，

---

1) 只要它是连续并且逐段光滑的(第 636 页)，这是物理中最重要的情形。

它与周期外力具有相同的节奏. 从而我们试着去求这微分方程的一个形如

$$x = \sigma e^{i\omega t} \tag{9}$$

的解，这里我们仅需确定与时间无关的因子 σ. 如果我们把这个式子和它的导数 $\dot{x} = i\omega\sigma e^{i\omega t}$, $\ddot{x} = -\omega^2\sigma e^{i\omega t}$ 代入微分方程，并且消去公因子 $e^{i\omega t}$，就得到方程

$$-m\omega^2\sigma + ir\omega\sigma + k\sigma = c,$$

或

$$\sigma = \frac{c}{-m\omega^2 + ir\omega + k}. \tag{10}$$

反过来，我们看到，对于 σ 的这个值，表达式 $\sigma e^{i\omega t}$ 确是微分方程的一个解. 但是，为了清楚地说明这个结果的意义，我们必须进行一些变换. 首先，我们把复因子 σ 写成形式

$$\sigma = c\,\frac{k - m\omega^2 - ir\omega}{(k - m\omega^2)^2 + r^2\omega^2} = c\alpha e^{-i\omega\delta}, \tag{11}$$

其中正“畸变因子” α 和“位相位移” $\omega\delta$ 是依据给定的量 m, r, k, 由方程

$$\alpha^2 = \frac{1}{(k - m\omega^2)^2 + r^2\omega^2},$$

$$\sin\omega\delta = r\omega\alpha, \quad \cos\omega\delta = (k - m\omega^2)\alpha$$

来表示的. 使用这种记号我们的解取形式

$$x = c\alpha e^{i\omega(t-\delta)},$$

这个结果的意义如下： 力 $c\cos\omega t$ 所产生的相应“效果”为 $c\alpha\cos\omega(t-\delta)$，而力 $c\sin\omega t$ 所产生的相应效果为 $c\alpha\sin\omega(t-\delta)$.

因此我们看到效果与力是属于相同类型的函数，亦即，是一个非阻尼的振动. 这个振动不同于代表力的振动之处在于，振幅按比例 $\alpha:1$ 增加，而位相改变了 $\omega\delta$. 当然，不用复数符号也容易得到相同的结果，但是要做一些较长的计算.

根据本节开始的一般说明，依据找到的这一个解，我们就全部解决了问题；因为叠加上任何一个自由振动，我们就能够得到最普

遍的强迫振动.

把这些结果收集起来,我们叙述如下:

微分方程

$$m\ddot{x}+r\dot{x}+kx=ce^{i\omega t}$$

(其中 $r\neq 0$) 的完全积分是 $x=c\alpha e^{i\omega(t-\delta)}+u$, 其中 u 是相应的齐次方程 $m\ddot{x}+r\dot{x}+kx=0$ 的完全积分,而量 α 和 δ 是用方程

$$\alpha^2=\frac{1}{(k-m\omega^2)^2+r^2\omega^2},$$

$$\sin\omega\delta=r\omega\alpha,\qquad \cos\omega\delta=(k-m\omega^2)\alpha \tag{12}$$

来确定的.

这个一般解中的常数留给我们一种把解作得适合任何一个初始状态的可能性,也就是,对任意指定的值 x_0 和 $\dot{x}_0$ 都能够选取这些常数使得 $x(0)=x_0$ 和 $\dot{x}(0)=\dot{x}_0$.

c. 共振曲线

为了获得对我们已经得到的解和它在应用上的意义的理解,我们将研究作为"激发频率"ω 的函数的畸变因子 α, 亦即,函数

$$\phi(\omega)=\frac{1}{\sqrt{(k-m\omega^2)^2+r^2\omega^2}}. \tag{13}$$

这样一种仔细的研究是由于下述事实的推动,即,对于给定的常数 k, m, r, 或如我们所说的对于一个给定的"振动系统",我们可以认为这个系统受到圆频率很不相同的周期激发力的作用,而对这些广泛不同的激发力来考虑微分方程的解是重要的. 为了便于描述这个函数我们引入量 $\omega_0=\sqrt{k/m}$. 这个数 ω_0 是自由振动系统当摩擦力为零时所应具有的圆频率;或简短地说,它是无阻尼系统的自然频率(参看第 676 页). 由于有摩擦力 r, 自由系统的真实频率并不等于 ω_0, 而是

$$\nu=\sqrt{\frac{k}{m}-\frac{r^2}{4m^2}}.$$

这里我们假设 $4km - r^2 > 0$.（若不是这种情形，则自由系统没有频率；它是非周期的.）

当激发频率趋于无穷时，函数 $\phi(\omega)$ 逐渐趋于零值，而且事实上它以阶 $1/\omega^2$ 趋于零. 此外，$\phi(0) = 1/k$；换言之，一个频率为零和量为 1 的激发力，亦即，量为 1 的常力，使振动系统发生的位移等于 $1/k$. 在 ω 的正值区域内导数 $\phi'(\omega)$ 不可能为零，除非表达式 $(k - m\omega^2)^2 + r^2\omega^2$ 的导数为零，也就是，除非使

$$-4m\omega(k - m\omega^2) + 2r^2\omega = 0$$

成立的值 $\omega = \omega_1 > 0$. 为了使得这样一个值能够存在，显然，我们必须有 $2km - r^2 > 0$；在这种情况下

$$\omega_1 = \sqrt{\frac{k}{m} - \frac{r^2}{2m^2}} = \sqrt{\omega_0^2 - \frac{r^2}{2m^2}}.$$

因为函数 $\phi(\omega)$ 处处为正，对于 ω 的小值来说是单调增加的，并且在无穷远为零，所以这个值 ω_1 必定给出一个最大值. 我们称这个频率 ω_1 为系统的"共振频率".

把 ω_1 的表达式代入，我们找到最大值是

$$\phi(\omega_1) = \frac{1}{r\sqrt{(k/m - r^2/4m^2)}}.$$

当 $r \to 0$，这个值无限增大. 对于 $r = 0$，亦即，对于一个无阻尼振动系统，函数 $\phi(\omega)$ 在值 $\omega = \omega_1$ 处有一个无穷的间断. 这是一种极限情形，我们在后面将给以特殊的考虑.

函数 $\phi(\omega)$ 的图象称为系统的**共振曲线**. 当 $\omega = \omega_1$ 时（并且因此对于自然频率的邻域内小的值 r 来说）振幅 $\alpha = \phi(\omega)$ 的畸变特别大，这个事实是"共振现象"的数学表示，对于固定的值 m 和 k 来说，当 r 变得越来越小时它变得越来越明显.

在图 9.3 中，我们简略地画出了一族共振曲线，都对应于值 $m = 1$ 和 $k = 1$，因而对应于 $\omega_0 = 1$，但是具有不同的值 $D = \frac{1}{2}r$. 我们看到，对于小的 D 值在邻近 $\omega = 1$ 处很好地显示了共振；在极限情形 $D = 0$，$\phi(\omega)$ 在 $\omega = 1$ 处没有最大值而有一个

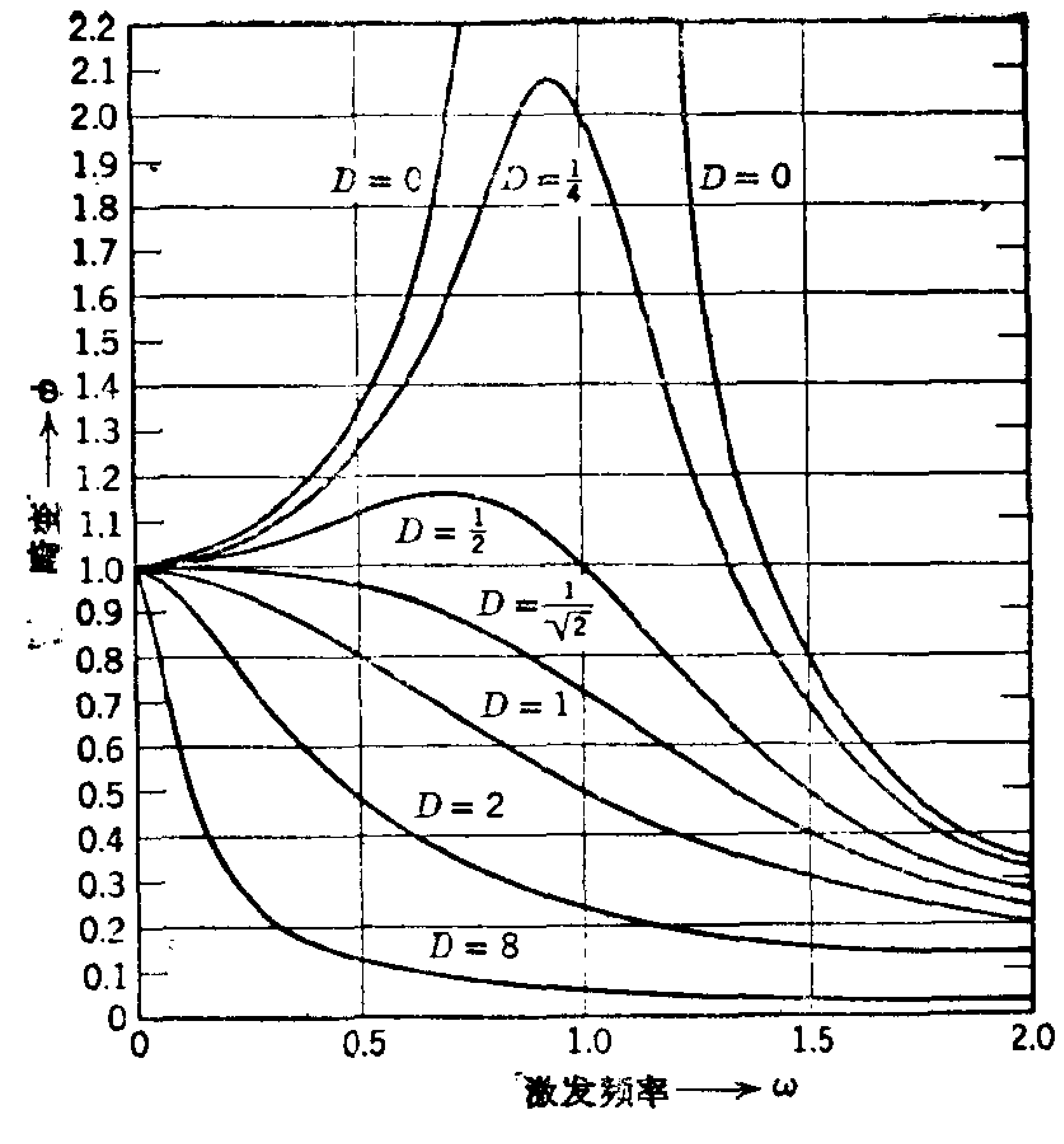


图 9.3　共振曲线

无穷的间断。当 D 增加时最大值向左移动，并且当值 $D = \frac{1}{\sqrt{2}}$ 时，我们有 $\omega_1 = 0$. 在最后这种情形，具有水平切线的点已经移动到原点，并且最大值已经消失. 如果 $D > 1/\sqrt{2}$ 则 $\phi'(\omega)$ 不为零；共振曲线不再有最大值，并且不再发生共振.

一般来说，条件

$$2km - r^2 \leqslant 0$$

一旦变成真实的时候共振现象就停止. 在等号成立的情形，共振曲线在 $\omega_1 = 0$ 处达到最大高度 $\phi(0) = 1/k$；在这里它的切线是水平的，并且切线在经历了几乎是水平线的初始过程之后它就向零倾斜.

d. 振动的进一步讨论

然而，我们不能满足于以上的讨论. 为了真正理解强迫运动的现象，还应当强调另外一点. 特殊积分 $c\alpha e^{i\omega(t-\delta)}$ 应看做是完全

积分

$$x(t) = c\alpha e^{i\omega(t-\delta)} + c_1u_1 + c_2u_2$$

的一种极限状态，这就是随着时间的前进完全积分越来越逼近于特殊积分；因为叠加在特殊积分上的自由振动 $c_1u_1 + c_2u_2$ 随着时间的流逝而消失．如果 r 小，消失将缓慢地发生，如果 r 大，消失将迅速地发生．

例如，我们假设运动开始时(亦即在时刻 $t = 0$ 时)这个系统是静止的，从而 $x(0)=0$ 和 $\dot{x}(0)=0$．从这里我们能够确定常数 c_1 和 c_2，并且立即看出它们两者不都为零．甚至当激发频率近似地或精确地等于 ω_1，从而共振发生时，最大的振幅 $\alpha = \phi(\omega_1)$ 在开始也不会出现．相反地，它将被函数 $c_1u_1 + c_2u_2$ 所掩盖，而当这个函数消失时，它才开始出现，并且，r 增长越小，它出现得越慢．

对于无阻尼系统（亦即对于 $r = 0$）当激发频率等于自然圆频率 $\omega_0 = \sqrt{k/m}$ 时，我们的解法失效，因为这时 $\phi(\omega_0)$ 是无穷．因而关于方程 $m\ddot{x} + kx = e^{i\omega t}$ 我们不能得到形式为 $\sigma e^{i\omega t}$ 的解．然而，我们能立即得到形式为 $x = \sigma te^{i\omega t}$ 的一个特解．如果我们把这个式子代入微分方程，并记着

$$\dot{x} = \sigma e^{i\omega t}(1 + i\omega t), \qquad \ddot{x} = \sigma e^{i\omega t}(2i\omega - t\omega^2),$$

我们就有

$$\sigma(2im\omega - m\omega^2 t + kt) = 1;$$

因而，由于 $m\omega^2 = k$，我们得到

$$\sigma = \frac{1}{2im\omega}.$$

这样，当共振发生在一个无阻尼系统里时，我们有一个解

$$x = \frac{t}{2im\omega}e^{i\omega t} = \frac{t}{2i\sqrt{km}}e^{i\omega t}.$$

使用实的记号，当 $f(t) = \cos\omega t$ 时，我们有

$$x = \frac{1}{2}\frac{t}{\sqrt{k/m}}\sin\omega t;$$

而当 $f(t)=\sin\omega t$ 时，我们有

$$x=-\frac{1}{2}\frac{1}{\sqrt{km}}\cos\omega t.$$

因此我们看到，我们已经找到一个函数，它可以看作是一个振动，但是这个振动的振幅随时间成比例地增大. 叠加上去的自由振动因为是无阻尼的，所以不会消失；但是它保留它原来的振幅，并且与特殊强迫振动的不断增大的振幅相比较变得不重要了. 在这种情形下，解在随时间的继续而不断增大的正负界限之间作来回的振动，这个事实表现了一个无阻尼系统共振函数的无穷间断点的实际意义.

e. 关于记录仪器构造的说明

在物理和工程的很多种应用当中，上面一小节的讨论是极为重要的. 联系很多仪器，如电流计，地震仪，无线电接收机中的振荡电路，以及传声器的膜片等，问题都是要记录由于一个周期外力的作用而产生的振动位移 x. 在这种情形下，量 x 满足我们的微分方程，至少对第一次近似是如此.

若 T 是周期外力的振动周期，则我们能把这个力展成一个富氏级数，具有如下的形式

$$f(t)=\sum_{l=-\infty}^{\infty}\gamma_l e^{il(2\pi/T)t},$$

或者说得更清楚些，我们能够认为在充分精确的程度以内把它表示成了仅由有限项组成的三角和 $\sum_{l=-N}^{N}\gamma_l e^{il(2\pi/T)t}$. 根据叠加原理（第 679 页），微分方程的解 $x(t)$，暂且不论叠加的自由振动，可以表示成一个无穷级数[1] 形如

$$x(t)=\sum_{l=-\infty}^{\infty}\sigma_l e^{il(2\pi/T)t}.$$

1）此处将不讨论收敛性问题.

或近似地表示成一个有限形式如

$$x(t)=\sum_{l=-N}^{N}\sigma_l e^{il(2\pi/T)t}.$$

根据前面的结果

$$\sigma_l=\gamma_l\alpha_l e^{-i\delta_l(2\pi l/T)}$$

和

$$\alpha_l^2=\frac{1}{\left(k-ml^2\frac{4\pi^2}{T^2}\right)^2+r^2l^2\frac{4\pi^2}{T^2}},$$

$$\tan\frac{2\pi l}{T}\delta_l=\frac{2\pi l r}{T\left(k-m\frac{4\pi^2l^2}{T^2}\right)}.$$

于是我们能够把一个任意周期外力所起的作用用下面的方式加以描述：如果我们把激发力分解成纯周期的分量，即傅立叶级数的单项，那么每个分量都引起它自己的振幅的畸变和位相位移，然后把分别的效果叠加起来就是总效果．如果我们仅对振幅畸变感兴趣（位相位移在应用中的重要性[1]仅是第二位的，而且也可以用振幅畸变同样的方法进行讨论），共振曲线的研究给了我们关于记录仪器反映激发外力的运动状态的完全信息．对于很大的值 l 或 $\omega[=2\pi/T]$，激发频率在位移 x 上的效果将是几乎不能觉察的．另一方面，在 ω_1 的邻域内的所有激发频率，即（圆）共振频率，都对于量 x 有显著影响．

在物理测量和记录仪器的构造中，常数 m，r 和 k 由我们自由选择，至少在一个很大的范围内是这样．这些常数应当选择得使共振曲线的形状尽可能适合于问题中的测量的特殊需要．这里有两种考虑起着支配作用．第一，我们希望仪器尽可能地灵敏；即对于问题中的所有频率 ω 来说 α 的值应尽可能地大．对于 ω 的小值来说，如我们已经看到的，α 是近似地与 $1/k$ 成比例，因而对

1) 例如，因为它对于人类的听觉是不可觉察的．

小的激发频率来说 $1/k$ 是仪器灵敏度的一个度量．因此灵敏度够通过 $1/k$ 的增大，亦即，通过恢复力的变弱来增大．

另一个要点是需要有抗畸变的相对自由度．设表达式 $f(t)=\sum_{l=-N}^{N}\gamma_l e^{il(2\pi/T)t}$ 足够近似于激发力．如果对于所有圆频率 $\omega\leqslant N(2\pi/T)$ 畸变因子都近似地具有相同的值，那么我们就说仪器带有抗畸变的相对自由度而记录着激发力 $f(t)$．这个条件是必不可少的，如果我们想要直接从仪器的性能来推导出关于激发过程的论断，例如，要使一个记录器或收音机要再产生与强度近似成比例的高低的音乐声调．要这种再生相对地“无畸变”的这种要求是永远不能精确地满足的，因为共振曲线上不存在严格的水平部分．然而我们能够试图用这样方式来选择仪器的常数 m,k,r，使得没有显著的共振发生，并且也用这样方式来选择使曲线在开始的时候具有一个水平切线，从而对于 ω 的小值来说，$\varphi(\omega)=\omega$ 保持近似于常数．如同上面我们已经学习过的，我们能够做到这点，只要设

$$2km-r^2=0.$$

给定常数 m 和常数 k 后，我们适当调整摩擦 r 就能满足这个要求，例如，在电路中插进适当选择的电阻．于是共振曲线指给我们，从频率 0 到接近于无阻尼系统的自然圆频率 ω_0 仪器几乎无畸变，而在这个频率以上阻尼是相当大的．因此，先把 m 选得很小，再把 k 选得很大使得无阻尼系统的自然圆频率 ω_0 大于任何一个在我们所考虑之下的激发圆频率，然后按照方程 $2km-r^2=0$ 选择阻尼因子 r，这样我们在一个给定的频率区间内就得到了抗畸变的相对自由度．